

**PALM INTRANET**

Day : Friday  
 Date: 7/8/2005  
 Time: 14:59:40

**Inventor Name Search Result**

Your Search was:

Last Name = BRUECKEN

First Name = CARL

Application#	Patent#	Status	Date Filed	Title	Inventor Name 8
<a href="#"><u>60301454</u></a>	Not Issued	159	06/29/2001	SYSTEM FOR NOTIFYING AN ONLINE CLIENT OF A MOBILE VENDOR	BRUECKEN, CARL JOSEPH
<a href="#"><u>60301453</u></a>	Not Issued	159	06/29/2001	DESCRIPTORS TO SUPPLEMENT ONLINE BANKING	BRUECKEN, CARL
<a href="#"><u>60301452</u></a>	Not Issued	159	06/29/2001	ELECTRONIC JOURNAL WITH LOCATION STAMP	BRUECKEN, CARL
<a href="#"><u>10914075</u></a>	Not Issued	030	08/10/2004	PASSIVE MONITORING OF USER INTERACTION WITH A BROWSER APPLICATION	BRUECKEN, CARL
<a href="#"><u>10134440</u></a>	Not Issued	030	04/30/2002	SYSTEM FOR NOTIFYING AN ONLINE CLIENT OF A MOBILE VENDOR	BRUECKEN, CARL
<a href="#"><u>09964453</u></a>	Not Issued	071	09/28/2001	AUTOMATED ELECTRONIC DICTIONARY	BRUECKEN, CARL
<a href="#"><u>09922000</u></a>	Not Issued	083	08/06/2001	ELECTRONIC JOURNAL WITH LOCATION STAMP	BRUECKEN, CARL
<a href="#"><u>09917676</u></a>	Not Issued	071	07/31/2001	METHOD FOR SUPPLEMENTING DESCRIPTORS FOR ONLINE BANKING TRANSACTION STATEMENTS	BRUECKEN, CARL

Inventor Search Completed: No Records to Display.

Search Another: Inventor	<input style="width: 100%; border: 1px solid black; height: 1.2em; font-size: 0.9em;" type="text" value="BRUECKEN"/>	<input style="width: 100%; border: 1px solid black; height: 1.2em; font-size: 0.9em;" type="text" value="CARL"/>
	<input style="border: 1px solid black; width: 60px; height: 1.2em; font-size: 0.8em; background-color: #e0e0e0;" type="button" value="Search"/>	

To go back use Back button on your browser toolbar.

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | Home page

 **PORTAL**  
USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)  
**Search:**  The ACM Digital Library  The Guide  
 + "electronic dictionary" and + "context"

**THE ACM DIGITAL LIBRARY**

 [Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Terms used electronic dictionary and context

09/964, 453

Found 58 of 157,956

Sort results by

 [Save results to a Binder](#)[Try an Advanced Search](#)

Display results

 [Search Tips](#)[Try this search in The ACM Guide](#) Open results in a new window

Results 1 - 20 of 58

Result page: **1** [2](#) [3](#) [next](#)Relevance scale **1 Placing search in context: the concept revisited**January 2002 **ACM Transactions on Information Systems (TOIS)**, Volume 20 Issue 1Full text available:  [pdf\(926.20 KB\)](#)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Keyword-based search engines are in widespread use today as a popular means for Web-based information retrieval. Although such systems seem deceptively simple, a considerable amount of skill is required in order to satisfy non-trivial information needs. This paper presents a new conceptual paradigm for performing search in context, that largely automates the search process, providing even non-professional users with highly relevant results. This paradigm is implemented in practice in the Intelli ...

**Keywords:** Search, context, invisible web, semantic processing, statistical natural language processing

**2 Placing search in context: the concept revisited**

Lev Finkelstein, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, Eytan Ruppin

April 2001 **Proceedings of the 10th international conference on World Wide Web**Full text available:  [pdf\(235.96 KB\)](#)Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

**Keywords:** context, invisible web, search, semantic processing, statistical natural language processing

**3 The textual development of non-stereotypic concepts**

Karin Haenelt, Michael Könyves-Tóth

April 1991 **Proceedings of the fifth conference on European chapter of the Association for Computational Linguistics**Full text available:  [pdf\(643.80 KB\)](#)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#) [Publisher Site](#)

In this paper the text theoretical foundation of our text analysis system KONTEXT is described. The basic premise of the KONTEXT model is that new concepts are communicated by using the mechanisms of text constitution. The text model used assumes that the

information conveyed in a text and the information describing its contextual organization can be structured into five layers (sentence structure, information on thematic progression, referential structure, conceptual representation of the text ...)

4 [Japanese OCR error correction using character shape similarity and statistical language model](#)

Masaaki Nagata  
August 1998

Full text available:  [pdf\(686.18 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)

We present a novel OCR error correction method for languages without word delimiters that have a large character set, such as Japanese and Chinese. It consists of a statistical OCR model, an approximate word matching method using character shape similarity, and a word segmentation algorithm using a statistical language model. By using a statistical OCR model and character shape similarity, the proposed error corrector outperforms the previously published method. When the baseline character recog ...

5 [Content-based retrieval: VideoQA: question answering on news video](#)

Hui Yang, Lekha Chaisorn, Yunlong Zhao, Shi-Yong Neo, Tat-Seng Chua  
November 2003 **Proceedings of the eleventh ACM international conference on Multimedia**

Full text available:  [pdf\(592.26 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

When querying a news video archive, the users are interested in retrieving precise answers in the form of a summary that best answers the query. However, current video retrieval systems, including the search engines on the web, are designed to retrieve documents instead of precise answers. This research explores the use of question answering (QA) techniques to support personalized news video retrieval. Users interact with our system, VideoQA, using short natural language questions with implicit ...

**Keywords:** transcript error correction, video question answering, video retrieval, video summarization

6 [Recognition of the coherence relation between te-linked clauses](#)

Akira Oishi, Yuji Matsumoto  
August 1998

Full text available:  [pdf\(643.55 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)  
 [Publisher Site](#)

This paper describes a method for recognizing coherence relations between clauses which are linked by *te* in Japanese—a translational equivalent of English *and*. We consider that the coherence relations are categories each of which has a prototype structure as well as the relationships among them. By utilizing this organization of the relations, we can infer an appropriate relation from the semantic structures of the clauses between which that relation holds. We carried out an experi ...

7 [The FINITE STRING newsletter: Abstracts of current literature](#)

Computational Linguistics Staff  
July 1984 **Computational Linguistics**, Volume 10 Issue 3-4

Full text available:  [pdf\(2.30 MB\)](#) Additional Information: [full citation](#)  
 [Publisher Site](#)

8 A part of speech estimation method for Japanese unknown words using a statistical model of morphology and context

Masaaki Nagata

June 1999 **Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics**

Full text available:  pdf(765.00 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

We present a statistical model of Japanese unknown words consisting of a set of length and spelling models classified by the character types that constitute a word. The point is quite simple: different character sets should be treated differently and the changes between character types are very important because Japanese script has both ideograms like Chinese <i>(kanji)</i> and phonograms like English <i>(katakana)</i>. Both word segmentation accuracy and part of speech taggin ...

9 Lexicon: Interlingual lexical organisation for multilingual lexical databases in NADIA

Gilles Sérasset

August 1994 **Proceedings of the 15th conference on Computational linguistics - Volume 1**

Full text available:  pdf(442.59 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

We propose a lexical organisation for multilingual lexical databases (MLDB). This organisation is based on acceptions (word-senses). We detail this lexical organisation and show a mock-up built to experiment with it. We also present our current work in defining and prototyping a specialised system for the management of accepton-based MLDB.

**Keywords:** acception, linguistic structure, multilingual lexical database

10 CYC, WordNet, and EDR: critiques and responses

Doug Lenat, George Miller, Toshio Yokoi

November 1995 **Communications of the ACM**, Volume 38 Issue 11

Full text available:  pdf(106.09 KB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

I applaud Miller's WordNet project and feel that there is much in common in our approaches, even though there are fundamental differences in the two expressions of that spirit. Here, I list the four differences I noted, closing with a crucial observation concerning the common spirit in our work.

11 Structural analysis of cooking preparation steps in Japanese

Reiko Hamada, Ichiro Ide, Shuichi Sakai, Hidehiko Tanaka

November 2000 **Proceedings of the fifth international workshop on on Information retrieval with Asian languages**

Full text available:  pdf(769.13 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

We propose a method to create process flow graphs automatically from textbooks for cooking programs. This is realized by understanding context by narrowing down the domain to cooking, and making use of domain specific constraints and knowledge. Since it is relatively easy to extract significant keywords from cooking procedures, we create a domain specific dictionary by statistical methods, and propose a structural analysis method using the dictionary. In order to evaluate the ability of the p ...

**Keywords:** cookbooks, domain specific dictionary, preparation steps, structural analysis

12 Posters: abstracts

Elizabeth D. Liddy

August 1996 **Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available:  pdf(288.41 KB) Additional Information: [full citation](#), [index terms](#)

**13 Papers: Extraction of lexical translations from non-aligned corpora** 

Kumiko Tanaka, Hideya Iwasaki

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 2**

Full text available:  pdf(571.08 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

A method for extracting lexical translations from non-aligned corpora is proposed to cope with the unavailability of large aligned corpus. The assumption that "translations of two co-occurring words in a source language also co-occur in the target language" is adopted and represented in the stochastic matrix formulation. The translation matrix provides the co-occurring information translated from the source into the target. This translated co-occurring information should resemble that of the ori ...

**14 Using decision trees to construct a practical parser** 

Masahiko Haruno, Satoshi Shirai, Yoshifumi Ooyama

August 1998

Full text available:  pdf(635.46 KB)  Publisher Site Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This paper describes novel and practical Japanese parsers that uses decision trees. First, we construct a single decision tree to estimate modification probabilities; how one phrase tends to modify another. Next, we introduce a boosting algorithm in which several decision trees are constructed and then combined for probability estimation. The two constructed parsers are evaluated by using the EDR Japanese annotated corpus. The single-tree method outperforms the conventional Japanese stochastic m ...

**15 Application of OODB and SGML techniques in text database: an electronic dictionary system** 

Jian Zhang

March 1995 **ACM SIGMOD Record**, Volume 24 Issue 1

Full text available:  pdf(557.23 KB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

An electronic dictionary system (EDS) is developed with object-oriented database techniques based on ObjectStore. The EDS is composed of two parts: the Database Building Program (DBP), and the Database Querying Program (DQP). DBP reads in a dictionary encoded in SGML tags, and builds a database composed of a collection of trees which holds dictionary entries, and several lists which contain items of various lexical categories. With text exchangeability introduced by the SGML, DBP is able to acco ...

**Keywords:** SGML, object-oriented databases, text database

**16 Large-scale resources: The automatic creation of lexical entries for a multilingual MT system** 

David Farwell, Louise Guthrie, Yorick Wilks

August 1992 **Proceedings of the 14th conference on Computational linguistics - Volume 2**

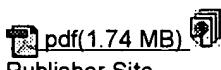
Full text available:  pdf(436.57 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

In this paper, we describe a method of extracting information from an on-line resource for the construction of lexical entries for a multi-lingual, interlingual MT system (ULTRA). We have been able to automatically generate lexical entries for interlingual concepts corresponding to nouns, verbs, adjectives and adverbs. Although several features of these entries continue to be supplied manually we have greatly decreased the time required to generate each entry and see this as a promising method f ...

**17 Selective sampling for example-based word sense disambiguation**

Atsushi Fujii, Takenobu Tokunaga, Kentaro Inui, Hozumi Tanaka  
December 1998 **Computational Linguistics**, Volume 24 Issue 4

Full text available:



Additional Information: [full citation](#), [abstract](#), [references](#)

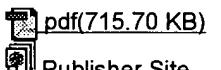
[Publisher Site](#)

This paper proposes an efficient example sampling method for example-based word sense disambiguation systems. To construct a database of practical size, a considerable overhead for manual sense disambiguation (overhead for supervision) is required. In addition, the time complexity of searching a large-sized database poses a considerable problem (overhead for search). To counter these problems, our method selectively samples a smaller-sized effective subset from a given example set for use in wor ...

**18 General-to-specific model selection for subcategorization preference**

Takehito Utsuro, Takashi Miyata, Yuji Matsumoto  
August 1998

Full text available:



Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

[Publisher Site](#)

This paper proposes a novel method for learning probability models of subcategorization preference of verbs. We consider the issues of *case dependencies* and *noun class generalization* in a uniform way by employing the maximum entropy modeling method. We also propose a new model selection algorithm which starts from the most *general* model and gradually examines more *specific* models. In the experimental evaluation, it is shown that both of the case dependencies and speci ...

**19 Lexicon: Analysis of scene identification ability of associative memory with pictorial dictionary**

Tatsuhiko Tsunoda, Hidehiko Tanaka  
August 1994 **Proceedings of the 15th conference on Computational linguistics - Volume 1**

Full text available:



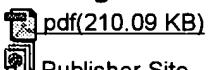
Additional Information: [full citation](#), [abstract](#), [references](#)

Semantic disambiguation depends on a process of defining the appropriate knowledge context. Recent research directions suggest a connectionist approach which use dictionaries, but there remain problems of scale, analysis, and interpretation. Here we focus on word disambiguation as scene selection, based on the Oxford Pictorial English Dictionary. We present a results of a spatial-scene identification ability using our original associative memory, We show both theoretical and experimental analysi ...

**20 Poster: Practical world modeling for NLP applications**

Lynn Carlson, Sergei Nirenburg  
March 1992 **Proceedings of the third conference on Applied natural language processing**

Full text available:



Additional Information: [full citation](#), [references](#)

[Publisher Site](#)

Results 1 - 20 of 58

Result page: [1](#) [2](#) [3](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2005 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

 **PORTAL**  
USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

**Search:**  The ACM Digital Library  The Guide

+ "electronic dictionary" and + "context"

 [Feedback](#) [Report a problem](#) [Satisfaction survey](#)

**THE ACM DIGITAL LIBRARY**Terms used [electronic dictionary](#) and [context](#)

Found 58 of 157,956

Sort results by [relevance](#) [Save results to a Binder](#)[Try an Advanced Search](#)Display results [expanded form](#) [Search Tips](#)[Try this search in The ACM Guide](#) [Open results in a new window](#)

Results 21 - 40 of 58

Result page: [previous](#) **1** [2](#) [3](#) [next](#)Relevance scale **21 A stochastic language model using dependency and its improvement by word clustering**Shinsuke Mori, Makoto Nagao  
August 1998Full text available:  [pdf\(550.06 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

In this paper, we present a stochastic language model for Japanese using dependency. The prediction unit in this model is an attribute of "bunsetsu". This is represented by the product of the head of content words and that of function words. The relation between the attributes of "bunsetsu" is ruled by a context-free grammar. The word sequences are predicted from the attribute using word n-gram model. The spell of Unknown word is predicted using character n-gram model. This model is robust in tha ...

**22 Algorithms for grapheme-phoneme translation for English and French: applications for database searches and speech synthesis**Michel Divay, Anthony J. Vitale  
December 1997 **Computational Linguistics**, Volume 23 Issue 4Full text available:   [pdf\(1.92 MB\)](#)  Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)  
[Publisher Site](#)

Letter-to-sound rules, also known as grapheme-to-phoneme rules, are important computational tools and have been used for a variety of purposes including word or name lookups for database searches and speech synthesis. These rules are especially useful when integrated into database searches on names and addresses, since they can complement orthographic search algorithms that make use of permutation, deletion, and insertion by allowing for a comparison with the phonetic equivalent. In databases, ph ...

**23 Automatic extraction of aspectual information from a monolingual corpus**Akira Oishi, Yuji Matsumoto  
July 1997Full text available:   [pdf\(712.44 KB\)](#)  Additional Information: [full citation](#), [abstract](#), [references](#)  
[Publisher Site](#)

This paper describes an approach to extract the aspectual information of Japanese verb phrases from a monolingual corpus. We classify verbs into six categories by means of the aspectual features which are defined on the basis of the possibility of co-occurrence with aspectual forms and adverbs. A unique category could be identified for 96% of the target

verbs. To evaluate the result of the experiment, we examined the meaning of *-teiru* which is one of the most fundamental aspectual markers ...

#### **24 Poster Sessions: A tagger/lemmatiser for Dutch medical language**

Peter Spyns

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 2**

Full text available:  pdf(362.08 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

In this paper, we want to describe a tagger/lemmatiser for Dutch medical vocabulary, which consists of a full-form dictionary and a morphological recogniser for unknown vocabulary coupled to an expert system-like disambiguation module. Attention is also paid to the main datastructures: a lexical database and feature bundles implemented as directed acyclic graphs. Some evaluation results are presented as well. The tagger/lemmatiser currently functions as a lexical front-end for a syntactic parser ...

#### **25 Natural language processing**

Yorick Wilks

January 1996 **Communications of the ACM**, Volume 39 Issue 1

Full text available:  pdf(238.26 KB) Additional Information: [full citation](#), [index terms](#)

#### **26 A hybrid Japanese parser with hand-crafted grammar and statistics**

Hiroshi Kanayama, Kentaro Torisawa, Yutaka Mitsuishi, Jun'ichi Tsujii

July 2000 **Proceedings of the 18th conference on Computational linguistics - Volume 1**

Full text available:  pdf(680.76 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This paper describes a hybrid parsing method for Japanese which uses both a hand-crafted grammar and a statistical technique. The key feature of our system is that in order to estimate likelihood for a parse tree, the system uses information taken from alternative partial parse trees generated by the grammar. This utilization of alternative trees enables us to construct a new statistical model called *Triplet/Quadruplet Model*. We show that this model can capture a certain tendency in Japan ...

#### **27 Part-of-speech induction from scratch**

Hinrich Schütze

June 1993 **Proceedings of the 31st annual meeting on Association for Computational Linguistics**

Full text available:  pdf(717.90 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

 Publisher Site

This paper presents a method for inducing the parts of speech of a language and part-of-speech labels for individual words from a large text corpus. Vector representations for the part-of-speech of a word are formed from entries of its near lexical neighbors. A dimensionality reduction creates a space representing the syntactic categories of unambiguous words. A neural net trained on these spatial representations classifies individual contexts of occurrence of ambiguous words. The method classif ...

#### **28 Technique for automatically correcting words in text**

Karen Kukich

December 1992 **ACM Computing Surveys (CSUR)**, Volume 24 Issue 4

Full text available:  pdf(6.23 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Research aimed at correcting words in text has focused on three progressively more difficult



Full text available:  pdf(605.73 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

This paper introduces a new approach to morpho-syntactic analysis through Humor 99 (<u>H</u>igh-speed <u>U</u>nification <u>Mor</u>phology), a reversible and unification-based morphological analyzer which has already been integrated with a variety of industrial applications. Humor 99 successfully copes with problems of agglutinative (e.g. Hungarian, Turkish, Estonian) and other (highly) inflectional languages (e.g. Polish, Czech, German) very effectively. T ...

**33 Special issue on using large corpora: I: Introduction to the special issue on computational linguistics using large corpora** 

Kenneth W. Church, Robert L. Mercer

March 1993 **Computational Linguistics**, Volume 19 Issue 1

Full text available:  pdf(1.80 MB)  Additional Information: [full citation](#), [references](#), [citations](#)  
[Publisher Site](#)

**34 Poster papers: Discovering word senses from text** 

Patrick Pantel, Dekang Lin

July 2002 **Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  pdf(661.99 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Inventories of manually compiled dictionaries usually serve as a source for word senses. However, they often include many rare senses while missing corpus/domain-specific senses. We present a clustering algorithm called CBC (Clustering By Committee) that automatically discovers word senses from text. It initially discovers a set of tight clusters called committees that are well scattered in the similarity space. The centroid of the members of a committee is used as the feature vector of the clus ...

**Keywords:** clustering, evaluation, machine learning, word sense discovery

**35 Poster Sessions: Word extraction from corpora and its part-of-speech estimation using distributional analysis** 

Shinsuke Mori, Makoto Nagao

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 2**

Full text available:  pdf(360.98 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Unknown words are inevitable at any step of analysis in natural language processing. We propose a method to extract words from a corpus and estimate the probability that each word belongs to given parts of speech (POSs), using a distributional analysis. Our experiments have shown that this method is effective for inferring the POS of unknown words.

**36 Lexicon: Noun phrasal entries in the EDR English word dictionary** 

A. Koizumi, M. Arioka, C. Harada, M. Sugimoto, L. Guthrie, C. Watts, R. Catizone, Y. Wilks

August 1994 **Proceedings of the 15th conference on Computational linguistics - Volume 1**

Full text available:  pdf(480.83 KB) Additional Information: [full citation](#), [references](#)

**Keywords:** lexicon Construction, resources for CL, universal features

## 37 Query term disambiguation for Web cross-language information retrieval using a search engine

Akira Maeda, Fatiha Sadat, Masatoshi Yoshikawa, Shunsuke Uemura

## **November 2000 Proceedings of the fifth international workshop on Information retrieval with Asian languages**

**Full text available:**  [pdf\(736.31 KB\)](#) **Additional Information:** [full citation](#), [abstract](#), [references](#), [citations](#)

With the worldwide growth of the Internet, research on Cross-Language Information Retrieval (CLIR) is being paid much attention. Existing CLIR approaches based on query translation require parallel corpora or comparable corpora for the disambiguation of translated query terms. However, those natural language resources are not readily available. In this paper, we propose a disambiguation method for dictionary-based query translation that is independent of the availability of such scarce langua ...

**Keywords:** WWW, cross-language information retrieval, mutual information, search engine

**38 Papers:** Inherited Feature-based Similarity Measure based on large semantic hierarchy and large text corpus

Hideki Hirakawa, Zhonghui Xu, Kenneth Haase

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 1**

Full text available: [pdf\(623.25 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

We describe a similarity calculation model called IFSM (Inherited Feature Similarity Measure) between objects (words/concepts) based on their common and distinctive features. We propose an implementation method for obtaining features based on abstracted triples extracted from a large text corpus utilizing taxonomical knowledge. This model represents an integration of traditional methods, i.e., relation based similarity measure and distribution based similarity measure. An experiment, using our n ...

39 A rule-based hyphenator for Modern Greek

Theodora I. Noussia

September 1997 **Computational Linguistics**, Volume 23 Issue 3

Full text available:  [pdf\(1.16 MB\)](#)  Additional Information: [full citation](#), [abstract](#), [references](#)  
[Publisher Site](#)

The purpose of this paper is to formally examine hyphenation as it pertains to Modern Greek with the aim of achieving accurate and thorough machine hyphenation. Grammar rules are interpreted and formally expressed in terms of regular expressions of word substrings, and exact hyphenation rules are derived. Vowel splitting, which traditionally is indicated in terms of prohibitive rather than explicit grammar rules, is examined in detail. Many ambiguities caused by circular definitions of the prohi ...

## 40 Large-scale resources: A Chinese corpus for linguistic research

Chu-Ren Huang, Keh-jiann Chen

August 1992 Proceedings of the 14th conference on Computational linguistics - Volume 4

**Full text available:**  pdf(259.12 KB) **Additional Information:** full citation, abstract, references, citations

This is a project note on the first stage of the construction of a comprehensive corpus of both Modern and Classical Chinese. The corpus is built with the dual aim of serving as the central database for Chinese language processing and for supporting in-depth linguistic research in Mandarin Chinese.

Results 21 - 40 of 58

Result page: [previous](#) [1](#) [2](#) [3](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2005 ACM, Inc.  
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)





**49 Subverting Structure: Data-Driven Diagram Generation**

Gene Golovchinsky, Klaus Reichenberger, Thomas Kamps

October 1995 **Proceedings of the 6th conference on Visualization '95**Full text available:  [pdf\(848.62 KB\)](#)[Publisher Site](#)Additional Information: [full citation](#), [abstract](#)

Diagrams are data representations that convey information predominantly through combinations of graphical elements rather than through other channels such as text or interaction. We have implemented a prototype called AVE (Automatic Visualization Environment) that generates diagrams automatically based on a generative theory of diagram design. According to this theory, diagrams are constructed based on the data to be visualized rather than by selection from a predefined set of diagrams. This app ...

**50 Computational methods ("paradigms"): The typology of unknown words: an experimental study of two corpora**

Xiaobo Ren, François Perrault

August 1992 **Proceedings of the 14th conference on Computational linguistics - Volume 1**Full text available:  [pdf\(464.19 KB\)](#) Additional Information: [full citation](#), [references](#)**51 Performance systems technology (PST) and computer-based instruction (CBI): tools for instructional designers in the 21st Century, part II**

Gloria A. Reece

November 1997 **ACM SIGDOC Asterisk Journal of Computer Documentation, Volume 21 Issue 4**Full text available:  [pdf\(304.72 KB\)](#) Additional Information: [full citation](#), [index terms](#)**52 Panel 4: The Internet a "natural" channel for language learning**

Inui Kentaro

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 2**Full text available:  [pdf\(126.67 KB\)](#) Additional Information: [full citation](#)**53 Short Papers: Ontology modeling tool with concept dictionary**

Yoichi Hiramatsu, Seiji Koide

January 2004 **Proceedings of the 9th international conference on Intelligent user interface**Full text available:  [pdf\(196.04 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

The usefulness of ontology is strongly dependent on the knowledge representation policy and its maintenance. The subject of knowledge representation and modeling tool has been one of the exciting themes among ontology scientists. Some ontology editing tools were born and grew up in the field of expert system and others designed originally by ontology research groups. Key features of the newly implemented tool are: reference to the concept dictionary to find out semantics of the words, and use of ...

**Keywords:** concept dictionary, editing tool, inference, ontology modeling, web service

**54 XRel: a path-based approach to storage and retrieval of XML documents using**

**relational databases**August 2001 **ACM Transactions on Internet Technology (TOIT)**, Volume 1 Issue 1Full text available:  pdf(264.27 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

This article describes XRel, a novel approach for storage and retrieval of XML documents using relational databases. In this approach, an XML document is decomposed into nodes on the basis of its tree structure and stored in relational tables according to the node type, with path information from the root to each node. XRel enables us to store XML documents using a fixed relational schema without any information about DTDs and also to utilize indices such as the B+

**Keywords:** XML query, XPath, text markup, text tagging

**55 New horizons in commercial and industrial AI**

Toshinori Munakata

November 1995 **Communications of the ACM**, Volume 38 Issue 11Full text available:  pdf(400.28 KB)Additional Information: [full citation](#), [abstract](#), [index terms](#)

AI as a field has undergone rapid growth in diversification and practicality. For the past 10 years, the repertoire of AI techniques has evolved and expanded. Scores of newer fields have recently been added to the traditional domains of practical AI. Although much practical AI is still best characterized as advanced computing rather than intelligence, applications in everyday commercial and industrial settings have certainly increased, especially since 1990. Additionally, A ...

**56 Cross-language Information Retrieval: Comparing cross-language query expansion techniques by degrading translation resources**

Paul McNamee, James Mayfield

August 2002 **Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval**Full text available:  pdf(267.21 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The quality of translation resources is arguably the most important factor affecting the performance of a cross-language information retrieval system. While many investigations have explored the use of query expansion techniques to combat errors induced by translation, no study has yet examined the effectiveness of these techniques across resources of varying quality. This paper presents results using parallel corpora and bilingual wordlists that have been deliberately degraded prior to query tr ...

**Keywords:** cross-language information retrieval, query expansion, query translation, translation resources

**57 Anchor text mining for translation of Web queries: A transitive translation approach**

Wen-Hsiang Lu, Lee-Feng Chien, Hsi-Jian Lee

April 2004 **ACM Transactions on Information Systems (TOIS)**, Volume 22 Issue 2Full text available:  pdf(280.55 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

To discover translation knowledge in diverse data resources on the Web, this article proposes an effective approach to finding translation equivalents of query terms and constructing multilingual lexicons through the mining of Web anchor texts and link structures. Although Web anchor texts are wide-scoped hypertext resources, not every particular pair of languages contains sufficient anchor texts for effective extraction of translations for Web queries. For more generalized applications, the app ...

**Keywords:** Multilingual translation, anchor text mining, competitive linking algorithm, cross-language Web search, cross-language information retrieval

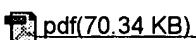
**58** [Electronic component information exchange \(ECIX\)](#)



Donald R. Cottrell

June 1997 **Proceedings of the 34th annual conference on Design automation - Volume 00**

Full text available:



[pdf\(70.34 KB\)](#)



Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

[Publisher Site](#)

A number of industry trends are shaping the requirements for IC and electronic equipment design. The density and complexity of circuit technologies have increased to a point where design cannot be performed without EDA tools. The availability of completely designed and verified reusable design components has become a major impediment to meeting required design productivity goals. Design reuse is moving down the package hierarchy to include chip design in addition to PCA design. At the same time, the wid ...

Results 41 - 58 of 58

Result page: [previous](#) [1](#) [2](#) [3](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2005 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:



[Adobe Acrobat](#)



[QuickTime](#)



[Windows Media Player](#)



[Real Player](#)

**PALM INTRANET**

Day : Friday  
 Date: 7/8/2005  
 Time: 11:49:22

**Application Number Information**Application Number: **09/964453****Assignments**Filing or 371(c) Date: **09/28/2001**Effective Date: **09/28/2001**Application Received: **09/28/2001**

Patent Number:

Issue Date: **00/00/0000**Date of Abandonment: **00/00/0000**Attorney Docket Number: **06975-202001****Status: 71 /RESPONSE TO NON-FINAL OFFICE ACTION ENTERED AND FORWARDED TO EXAMINER**Confirmation Number: **1659**Examiner Number: **78953 /WONG, LESLIE**Group Art Unit: **2167****IFW IMAGE**Class/Subclass: **707/001.000**Lost Case: **NO**

Interference Number:

Unmatched Petition: **NO**L&R Code: Secrecy Code: **1**Third Level Review: **NO**Secrecy Order: **NO**Status Date: **06/10/2005**Oral Hearing: **NO**Title of Invention: **AUTOMATED ELECTRONIC DICTIONARY**

Bar Code	PALM Location	Location Date	Charge to Loc	Charge to Name	Employee Name	Location
<b>09964453ZA</b>	<b>28C1</b>	<b>06/09/2005</b>	<b>No Charge to Location</b>	<b>No Charge to Name</b>	<b>NGUYEN,HUY</b>	<b>JEF/00/A 79</b>

**Appln Info** [Contents](#) [Petition Info](#) [Atty/Agent Info](#) [Continuity Data](#) [Foreign Data](#) [Inve](#)

**Search Another: Application#**

**or Patent#**

**PCT /**  **/**   **or PG PUBS #**

**Attorney Docket #**

**Bar Code #**

To go back use Back button on your browser toolbar.

Back to [PALM](#) | [ASSIGNMENT](#) | [OASIS](#) | Home page



[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

Search:  The ACM Digital Library  The Guide

+ "electronic dictionary" and + "context" and (subsequent <sen



## THE ACM DIGITAL LIBRARY

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Terms used **electronic dictionary** and  
**context** and **subsequent sentence access**

Found 58 of 157,956

Sort results  
by

[Save results to a Binder](#)

[Try an Advanced Search](#)

Display  
results

[Search Tips](#)

[Try this search in The ACM Guide](#)

[Open results in a new window](#)

Results 1 - 20 of 58

Result page: **1** [2](#) [3](#) [next](#)

Relevance scale

### **1 Technique for automatically correcting words in text**

Karen Kukich

December 1992 **ACM Computing Surveys (CSUR)**, Volume 24 Issue 4

Full text available: [pdf\(6.23 MB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Research aimed at correcting words in text has focused on three progressively more difficult problems: (1) nonword error detection; (2) isolated-word error correction; and (3) context-dependent word correction. In response to the first problem, efficient pattern-matching and n-gram analysis techniques have been developed for detecting strings that do not appear in a given word list. In response to the second problem, a variety of general and application-specific spelling cor ...

**Keywords:** n-gram analysis, Optical Character Recognition (OCR), context-dependent spelling correction, grammar checking, natural-language-processing models, neural net classifiers, spell checking, spelling error detection, spelling error patterns, statistical-language models, word recognition and correction

### **2 Placing search in context: the concept revisited**

Lev Finkelstein, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, Eytan Ruppin

April 2001 **Proceedings of the 10th international conference on World Wide Web**

Full text available: [pdf\(235.96 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



**Keywords:** context, invisible web, search, semantic processing, statistical natural language processing

### **3 Placing search in context: the concept revisited**

January 2002 **ACM Transactions on Information Systems (TOIS)**, Volume 20 Issue 1

Full text available: [pdf\(926.20 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)



Keyword-based search engines are in widespread use today as a popular means for Web-

based information retrieval. Although such systems seem deceptively simple, a considerable amount of skill is required in order to satisfy non-trivial information needs. This paper presents a new conceptual paradigm for performing search in context, that largely automates the search process, providing even non-professional users with highly relevant results. This paradigm is implemented in practice in the Intelli ...

**Keywords:** Search, context, invisible web, semantic processing, statistical natural language processing

**4 Content-based retrieval: VideoQA: question answering on news video**

Hui Yang, Lekha Chaisorn, Yunlong Zhao, Shi-Yong Neo, Tat-Seng Chua

November 2003 **Proceedings of the eleventh ACM international conference on Multimedia**

Full text available:  [pdf\(592.26 KB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

When querying a news video archive, the users are interested in retrieving precise answers in the form of a summary that best answers the query. However, current video retrieval systems, including the search engines on the web, are designed to retrieve documents instead of precise answers. This research explores the use of question answering (QA) techniques to support personalized news video retrieval. Users interact with our system, VideoQA, using short natural language questions with implicit ...

**Keywords:** transcript error correction, video question answering, video retrieval, video summarization

**5 Selective sampling for example-based word sense disambiguation**

Atsushi Fujii, Takenobu Tokunaga, Kentaro Inui, Hozumi Tanaka

December 1998 **Computational Linguistics**, Volume 24 Issue 4

Full text available:

 [pdf\(1.74 MB\)](#) 

Additional Information: [full citation](#), [abstract](#), [references](#)  
[Publisher Site](#)

This paper proposes an efficient example sampling method for example-based word sense disambiguation systems. To construct a database of practical size, a considerable overhead for manual sense disambiguation (overhead for supervision) is required. In addition, the time complexity of searching a large-sized database poses a considerable problem (overhead for search). To counter these problems, our method selectively samples a smaller-sized effective subset from a given example set for use in wor ...

**6 Algorithms for grapheme-phoneme translation for English and French: applications for database searches and speech synthesis**

Michel Divay, Anthony J. Vitale

December 1997 **Computational Linguistics**, Volume 23 Issue 4

Full text available:

 [pdf\(1.92 MB\)](#) 

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)  
[Publisher Site](#)

Letter-to-sound rules, also known as grapheme-to-phoneme rules, are important computational tools and have been used for a variety of purposes including word or name lookups for database searches and speech synthesis. These rules are especially useful when integrated into database searches on names and addresses, since they can complement orthographic search algorithms that make use of permutation, deletion, and insertion by allowing for a comparison with the phonetic equivalent. In databases, ph ...

**Poster Sessions: A tagger/lemmatiser for Dutch medical language**

Peter Spyns

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 2**

Full text available:  pdf(362.08 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

In this paper, we want to describe a tagger/lemmatiser for Dutch medical vocabulary, which consists of a full-form dictionary and a morphological recogniser for unknown vocabulary coupled to an expert system-like disambiguation module. Attention is also paid to the main datastructures: a lexical database and feature bundles implemented as directed acyclic graphs. Some evaluation results are presented as well. The tagger/lemmatiser currently functions as a lexical front-end for a syntactic parser ...

**8 Poster Sessions: Word extraction from corpora and its part-of-speech estimation using distributional analysis**

Shinsuke Mori, Makoto Nagao

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 2**

Full text available:  pdf(360.98 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Unknown words are inevitable at any step of analysis in natural language processing. We propose a method to extract words from a corpus and estimate the probability that each word belongs to given parts of speech (POSs), using a distributional analysis. Our experiments have shown that this method is effective for inferring the POS of unknown words.

**9 Cross-language Information Retrieval: Comparing cross-language query expansion techniques by degrading translation resources**

Paul McNamee, James Mayfield

August 2002 **Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available:  pdf(267.21 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The quality of translation resources is arguably the most important factor affecting the performance of a cross-language information retrieval system. While many investigations have explored the use of query expansion techniques to combat errors induced by translation, no study has yet examined the effectiveness of these techniques across resources of varying quality. This paper presents results using parallel corpora and bilingual wordlists that have been deliberately degraded prior to query tr ...

**Keywords:** cross-language information retrieval, query expansion, query translation, translation resources

**10 A part of speech estimation method for Japanese unknown words using a statistical model of morphology and context**

Masaaki Nagata

June 1999 **Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics**

Full text available:  pdf(765.00 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

We present a statistical model of Japanese unknown words consisting of a set of length and spelling models classified by the character types that constitute a word. The point is quite simple: different character sets should be treated differently and the changes between character types are very important because Japanese script has both ideograms like Chinese <i>(kanji)</i> and phonograms like English <i>(katakana)</i>. Both word segmentation

accuracy and part of speech taggin ...

**11 The textual development of non-stereotypic concepts**

Karin Haenelt, Michael Kónyves-Tóth

April 1991 **Proceedings of the fifth conference on European chapter of the Association for Computational Linguistics**

Full text available:  pdf(643.80 KB)

 Publisher Site

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

In this paper the text theoretical foundation of our text analysis system KONTEXT is described. The basic premise of the KONTEXT model is that new concepts are communicated by using the mechanisms of text constitution. The text model used assumes that the information conveyed in a text and the information describing its contextual organization can be structured into five layers (sentence structure, information on thematic progression, referential structure, conceptual representation of the text ...

**12 The FINITE STRING newsletter: Abstracts of current literature**

Computational Linguistics Staff

July 1984 **Computational Linguistics**, Volume 10 Issue 3-4

Full text available:  pdf(2.30 MB)

 Publisher Site

Additional Information: [full citation](#)

**13 Structural analysis of cooking preparation steps in Japanese**

Reiko Hamada, Ichiro Ide, Shuichi Sakai, Hidehiko Tanaka

November 2000 **Proceedings of the fifth international workshop on on Information retrieval with Asian languages**

Full text available:  pdf(769.13 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

We propose a method to create process flow graphs automatically from textbooks for cooking programs. This is realized by understanding context by narrowing down the domain to cooking, and making use of domain specific constraints and knowledge. Since it is relatively easy to extract significant keywords from cooking procedures, we create a domain specific dictionary by statistical methods, and propose a structural analysis method using the dictionary. In order to evaluate the ability of the p ...

**Keywords:** cookbooks, domain specific dictionary, preparation steps, structural analysis

**14 Special issue on using large corpora: I: Introduction to the special issue on computational linguistics using large corpora**

Kenneth W. Church, Robert L. Mercer

March 1993 **Computational Linguistics**, Volume 19 Issue 1

Full text available:

Additional Information: [full citation](#), [references](#), [citations](#)

[Publisher Site](#)

**15 A stochastic finite-state word-segmentation algorithm for Chinese**

Richard Sproat, William Gale, Chilin Shih, Nancy Chang

September 1996 **Computational Linguistics**, Volume 22 Issue 3

Full text available:

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

[Publisher Site](#)

The initial stage of text analysis for any NLP task usually involves the tokenization of the input into words. For languages like English one can assume, to a first approximation, that word boundaries are given by whitespace or punctuation. In various Asian languages, including Chinese, on the other hand, whitespace is never used to delimit words, so one must resort to lexical information to "reconstruct" the word-boundary information. In this paper we present a stochastic finite-state model whe ...

**16 Japanese OCR error correction using character shape similarity and statistical language model**

Masaaki Nagata  
August 1998

Full text available:  pdf(686.18 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

We present a novel OCR error correction method for languages without word delimiters that have a large character set, such as Japanese and Chinese. It consists of a statistical OCR model, an approximate word matching method using character shape similarity, and a word segmentation algorithm using a statistical language model. By using a statistical OCR model and character shape similarity, the proposed error corrector outperforms the previously published method. When the baseline character recog ...

**17 Using decision trees to construct a practical parser**

Masahiko Haruno, Satoshi Shirai, Yoshifumi Ooyama  
August 1998

Full text available:  pdf(635.46 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)  
 Publisher Site

This paper describes novel and practical Japanese parsers that uses decision trees. First, we construct a single decision tree to estimate modification probabilities; how one phrase tends to modify another. Next, we introduce a boosting algorithm in which several decision trees are constructed and then combined for probability estimation. The two constructed parsers are evaluated by using the EDR Japanese annotated corpus. The single-tree method outperforms the conventional Japanese stochastic m ...

**18 CYC, WordNet, and EDR: critiques and responses**

Doug Lenat, George Miller, Toshio Yokoi  
November 1995 **Communications of the ACM**, Volume 38 Issue 11

Full text available:  pdf(106.09 KB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

I applaud Miller's WordNet project and feel that there is much in common in our approaches, even though there are fundamental differences in the two expressions of that spirit. Here, I list the four differences I noted, closing with a crucial observation concerning the common spirit in our work.

**19 A stochastic language model using dependency and its improvement by word clustering**

Shinsuke Mori, Makoto Nagao  
August 1998

Full text available:  pdf(550.06 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

In this paper, we present a stochastic language model for Japanese using dependency. The prediction unit in this model is an attribute of "bunsetsu". This is represented by the product of the head of content words and that of function words. The relation between the attributes of "bunsetsu" is ruled by a context-free grammar. The word sequences are predicted from the attribute using word n-gram model. The spell of Unknow word is predicted using character n-gram model. This model is robust in tha ...

20 Automatic extraction of aspectual information from a monolingual corpus



Akira Oishi, Yuji Matsumoto  
July 1997

Full text available:  [pdf\(712.44 KB\)](#)

[Publisher Site](#)

Additional Information: [full citation](#), [abstract](#), [references](#)

This paper describes an approach to extract the aspectual information of Japanese verb phrases from a monolingual corpus. We classify verbs into six categories by means of the aspectual features which are defined on the basis of the possibility of co-occurrence with aspectual forms and adverbs. A unique category could be identified for 96% of the target verbs. To evaluate the result of the experiment, we examined the meaning of -teiru which is one of the most fundamental aspectual markers ...

Results 1 - 20 of 58

Result page: [1](#) [2](#) [3](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2005 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)



USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)**Search:**  The ACM Digital Library  The Guide

+"electronic dictionary" and +"context" and ((following or next



THE ACM DIGITAL LIBRARY

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Terms used **electronic dictionary** and  
**context** and **following or next sentence access** or **search**

Found 57 of 157,956

Sort results  
by  [Save results to a Binder](#)Display  
results  [Search Tips](#) [Open results in a new window](#)[Try an Advanced Search](#)[Try this search in The ACM Guide](#)

Results 1 - 20 of 57

Result page: **1** [2](#) [3](#) [next](#)

Relevance scale

**1** [Technique for automatically correcting words in text](#)

Karen Kukich

December 1992 **ACM Computing Surveys (CSUR)**, Volume 24 Issue 4Full text available: [pdf\(6.23 MB\)](#)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Research aimed at correcting words in text has focused on three progressively more difficult problems:(1) nonword error detection; (2) isolated-word error correction; and (3) context-dependent word correction. In response to the first problem, efficient pattern-matching and n-gram analysis techniques have been developed for detecting strings that do not appear in a given word list. In response to the second problem, a variety of general and application-specific spelling cor ...

**Keywords:** n-gram analysis, Optical Character Recognition (OCR), context-dependent spelling correction, grammar checking, natural-language-processing models, neural net classifiers, spell checking, spelling error detection, spelling error patterns, statistical-language models, word recognition and correction

Full text available: [!\[\]\(ec4bf86fbc20b4c99c0e88e3c82e29ee\_img.jpg\) pdf\(1.74 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)  
[Publisher Site](#)

This paper proposes an efficient example sampling method for example-based word sense disambiguation systems. To construct a database of practical size, a considerable overhead for manual sense disambiguation (overhead for supervision) is required. In addition, the time complexity of searching a large-sized database poses a considerable problem (overhead for search). To counter these problems, our method selectively samples a smaller-sized effective subset from a given example set for use in wor ...

**4 Content-based retrieval: VideoQA: question answering on news video**

Hui Yang, Lekha Chaisorn, Yunlong Zhao, Shi-Yong Neo, Tat-Seng Chua  
November 2003 **Proceedings of the eleventh ACM international conference on Multimedia**

Full text available: [!\[\]\(2de14ecdac8f3bd4221dec5cc1fcc44b\_img.jpg\) pdf\(592.26 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

When querying a news video archive, the users are interested in retrieving precise answers in the form of a summary that best answers the query. However, current video retrieval systems, including the search engines on the web, are designed to retrieve documents instead of precise answers. This research explores the use of question answering (QA) techniques to support personalized news video retrieval. Users interact with our system, VideoQA, using short natural language questions with implicit ...

**Keywords:** transcript error correction, video question answering, video retrieval, video summarization

**5 Placing search in context: the concept revisited**

January 2002 **ACM Transactions on Information Systems (TOIS)**, Volume 20 Issue 1

Full text available: [!\[\]\(71f9d84356bec8aef5a389e4c1b7f2e3\_img.jpg\) pdf\(926.20 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Keyword-based search engines are in widespread use today as a popular means for Web-based information retrieval. Although such systems seem deceptively simple, a considerable amount of skill is required in order to satisfy non-trivial information needs. This paper presents a new conceptual paradigm for performing search in context, that largely automates the search process, providing even non-professional users with highly relevant results. This paradigm is implemented in practice in the Intelli ...

**Keywords:** Search, context, invisible web, semantic processing, statistical natural language processing

**6 The FINITE STRING newsletter: Abstracts of current literature**

Computational Linguistics Staff  
July 1984 **Computational Linguistics**, Volume 10 Issue 3-4

Full text available: [!\[\]\(4f4d5875cccdb44c8f60eb9e86279722\_img.jpg\) pdf\(2.30 MB\)](#) Additional Information: [full citation](#)  
[Publisher Site](#)

**7 A stochastic finite-state word-segmentation algorithm for Chinese**

Richard Sproat, William Gale, Chilin Shih, Nancy Chang  
September 1996 **Computational Linguistics**, Volume 22 Issue 3

Full text available:

 pdf(1.91 MB) [Publisher Site](#)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

The initial stage of text analysis for any NLP task usually involves the tokenization of the input into words. For languages like English one can assume, to a first approximation, that word boundaries are given by whitespace or punctuation. In various Asian languages, including Chinese, on the other hand, whitespace is never used to delimit words, so one must resort to lexical information to "reconstruct" the word-boundary information. In this paper we present a stochastic finite-state model whe ...

## **8 Automatic extraction of aspectual information from a monolingual corpus**

Akira Oishi, Yuji Matsumoto  
July 1997

Full text available:  pdf(712.44 KB) [Publisher Site](#)Additional Information: [full citation](#), [abstract](#), [references](#)

This paper describes an approach to extract the aspectual information of Japanese verb phrases from a monolingual corpus. We classify verbs into six categories by means of the aspectual features which are defined on the basis of the possibility of co-occurrence with aspectual forms and adverbs. A unique category could be identified for 96% of the target verbs. To evaluate the result of the experiment, we examined the meaning of -teiru which is one of the most fundamental aspectual markers ...

## **9 Placing search in context: the concept revisited**

Lev Finkelstein, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, Eytan Ruppin

April 2001 **Proceedings of the 10th international conference on World Wide Web**Full text available:  pdf(235.96 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

**Keywords:** context, invisible web, search, semantic processing, statistical natural language processing

## **10 A part of speech estimation method for Japanese unknown words using a statistical model of morphology and context**

Masaaki Nagata

June 1999 **Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics**Full text available:  pdf(765.00 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

We present a statistical model of Japanese unknown words consisting of a set of length and spelling models classified by the character types that constitute a word. The point is quite simple: different character sets should be treated differently and the changes between character types are very important because Japanese script has both ideograms like Chinese <i>(kanji)</i> and phonograms like English <i>(katakana)</i>. Both word segmentation accuracy and part of speech taggin ...

## **11 Using decision trees to construct a practical parser**

Masahiko Haruno, Satoshi Shirai, Yoshifumi Ooyama  
August 1998

Full text available:  pdf(635.46 KB) [Publisher Site](#)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This paper describes novel and practical Japanese parsers that uses decision trees. First, we construct a single decision tree to estimate modification probabilities; how one phrase tends to modify another. Next, we introduce a boosting algorithm in which several decision trees are constructed and then combined for probability estimation. The two constructed parsers are evaluated by using the EDR Japanese annotated corpus. The single-tree method outperforms the conventional Japanese stochastic m ...

## **12 Structural analysis of cooking preparation steps in Japanese**

Reiko Hamada, Ichiro Ide, Shuichi Sakai, Hidehiko Tanaka

November 2000 **Proceedings of the fifth international workshop on on Information retrieval with Asian languages**

Full text available:  pdf(769.13 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

We propose a method to create process flow graphs automatically from textbooks for cooking programs. This is realized by understanding context by narrowing down the domain to cooking, and making use of domain specific constraints and knowledge. Since it is relatively easy to extract significant keywords from cooking procedures, we create a domain specific dictionary by statistical methods, and propose a structural analysis method using the dictionary. In order to evaluate the ability of the p ...

**Keywords:** cookbooks, domain specific dictionary, preparation steps, structural analysis

## **13 A stochastic language model using dependency and its improvement by word clustering**

Shinsuke Mori, Makoto Nagao

August 1998

Full text available:  pdf(550.06 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

In this paper, we present a stochastic language model for Japanese using dependency. The prediction unit in this model is an attribute of "bunsetsu". This is represented by the product of the head of content words and that of function words. The relation between the attributes of "bunsetsu" is ruled by a context-free grammar. The word sequences are predicted from the attribute using word n-gram model. The spell of Unknown word is predicted using character n-gram model. This model is robust in tha ...

## **14 Special issue on using large corpora: I: Introduction to the special issue on computational linguistics using large corpora**

Kenneth W. Church, Robert L. Mercer

March 1993 **Computational Linguistics**, Volume 19 Issue 1

Full text available:  pdf(1.80 MB)  Additional Information: [full citation](#), [references](#), [citations](#)  
[Publisher Site](#)

## **15 The textual development of non-stereotypic concepts**

Karin Haenelt, Michael Konyves-Tóth

April 1991 **Proceedings of the fifth conference on European chapter of the Association for Computational Linguistics**

Full text available:  pdf(643.80 KB)  Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)  
[Publisher Site](#)

In this paper the text theoretical foundation of our text analysis system KONTEXT is described. The basic premise of the KONTEXT model is that new concepts are communicated by using the mechanisms of text constitution. The text model used assumes that the

information conveyed in a text and the information describing its contextual organization can be structured into five layers (sentence structure, information on thematic progression, referential structure, conceptual representation of the text ...

#### **16 General-to-specific model selection for subcategorization preference**

Takehito Utsuro, Takashi Miyata, Yuji Matsumoto  
August 1998

Full text available:  pdf(715.70 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

 Publisher Site

This paper proposes a novel method for learning probability models of subcategorization preference of verbs. We consider the issues of *case dependencies* and *noun class generalization* in a uniform way by employing the maximum entropy modeling method. We also propose a new model selection algorithm which starts from the most *general* model and gradually examines more *specific* models. In the experimental evaluation, it is shown that both of the case dependencies and speci ...

#### **17 Lexicon: Analysis of scene identification ability of associative memory with pictorial dictionary**

Tatsuhiko Tsunoda, Hidehiko Tanaka  
August 1994 **Proceedings of the 15th conference on Computational linguistics - Volume 1**

Full text available:  pdf(538.99 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

Semantic disambiguation depends on a process of defining the appropriate knowledge context. Recent research directions suggest a connectionist approach which use dictionaries, but there remain problems of scale, analysis, and interpretation. Here we focus on word disambiguation as scene selection, based on the Oxford Pictorial English Dictionary. We present a results of a spatial-scene identification ability using our original associative memory, We show both theoretical and experimental analysi ...

#### **18 A hybrid Japanese parser with hand-crafted grammar and statistics**

Hiroshi Kanayama, Kentaro Torisawa, Yutaka Mitsuishi, Jun'ichi Tsuji  
July 2000 **Proceedings of the 18th conference on Computational linguistics - Volume 1**

Full text available:  pdf(680.76 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This paper describes a hybrid parsing method for Japanese which uses both a hand-crafted grammar and a statistical technique. The key feature of our system is that in order to estimate likelihood for a parse tree, the system uses information taken from alternative partial parse trees generated by the grammar. This utilization of alternative trees enables us to construct a new statistical model called *Triplet/Quadruplet Model*. We show that this model can capture a certain tendency in Japan ...

#### **19 Japanese OCR error correction using character shape similarity and statistical language model**

Masaaki Nagata  
August 1998

Full text available:  pdf(686.18 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

We present a novel OCR error correction method for languages without word delimiters that have a large character set, such as Japanese and Chinese. It consists of a statistical OCR model, an approximate word matching method using character shape similarity, and a word segmentation algorithm using a statistical language model. By using a statistical OCR model and character shape similarity, the proposed error corrector outperforms the previously published method. When the baseline character recog ...

20 [Learning probabilistic subcategorization preference by identifying case dependencies and optimal noun class generalization level](#)



Takehito Utsuro, Yuji Matsumoto

March 1997 **Proceedings of the fifth conference on Applied natural language processing**

Full text available:  [pdf\(844.41 KB\)](#)

[Publisher Site](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This paper proposes a novel method of learning probabilistic subcategorization preference. In the method, for the purpose of coping with the ambiguities of case dependencies and noun class generalization of argument/adjunct nouns, we introduce a data structure which represents a tuple of independent partial subcategorization frames. Each collocation of a verb and argument/adjunct nouns is assumed to be generated from one of the possible tuples of independent partial subcategorization frames. Par ...

Results 1 - 20 of 57

Result page: [1](#) [2](#) [3](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2005 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)



discovers word senses from text. It initially discovers a set of tight clusters called committees that are well scattered in the similarity space. The centroid of the members of a committee is used as the feature vector of the clus ...

**Keywords:** clustering, evaluation, machine learning, word sense discovery

## **24 Poster Sessions: Word extraction from corpora and its part-of-speech estimation using distributional analysis**

Shinsuke Mori, Makoto Nagao

August 1996 Proceedings of the 16th conference on Computational linguistics - Volume 2

Full text available:  pdf(360.98 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Unknown words are inevitable at any step of analysis in natural language processing. We propose a method to extract words from a corpus and estimate the probability that each word belongs to given parts of speech (POSs), using a distributional analysis. Our experiments have shown that this method is effective for inferring the POS of unknown words.

## **25 Large-scale resources: The automatic creation of lexical entries for a multilingual MT system**

David Farwell, Louise Guthrie, Yorick Wilks

August 1992 Proceedings of the 14th conference on Computational linguistics - Volume 2

**Full text available:**  pdf(436.57 KB) **Additional Information:** full citation, abstract, references, citations

In this paper, we describe a method of extracting information from an on-line resource for the construction of lexical entries for a multi-lingual, interlingual MT system (ULTRA). We have been able to automatically generate lexical entries for interlingual concepts corresponding to nouns, verbs, adjectives and adverbs. Although several features of these entries continue to be supplied manually we have greatly decreased the time required to generate each entry and see this as a promising method f ...

**26** Papers: Inherited Feature-based Similarity Measure based on large semantic hierarchy and large text corpus

Hideki Hirakawa, Zhonghui Xu, Kenneth Haase

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 1**

Full text available:  pdf(623.25 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

We describe a similarity calculation model called IFSM (Inherited Feature Similarity Measure) between objects (words/concepts) based on their common and distinctive features. We propose an implementation method for obtaining features based on abstracted triples extracted from a large text corpus utilizing taxonomical knowledge. This model represents an integration of traditional methods, i.e., relation based similarity measure and distribution based similarity measure. An experiment, using our n ...

## 27 Estimating understandability of software documents

Kari Laitinen

July 1996 ACM SIGSOFT Software Engineering Notes, Volume 21 Issue 4

**Full text available:**  pdf(878.94 KB) **Additional Information:** full citation, abstract, index terms

Software developers and maintainers need to read and understand source programs and other kinds of software documents in their work. Understandability of software documents is thus important. This paper introduces a method for estimating the understandability of

software documents. The method is based on a language theory according to which every software document is considered to contain a language of its own, which is a set of symbols. The understandability of documents written according to di ...

**28** [Poster: Practical world modeling for NLP applications](#)  
Lynn Carlson, Sergei Nirenburg  
March 1992 **Proceedings of the third conference on Applied natural language processing**  
Full text available:  [pdf\(210.09 KB\)](#) Additional Information: [full citation](#), [references](#)  
 [Publisher Site](#)

**29** [New horizons in commercial and industrial AI](#)  
Toshinori Munakata  
November 1995 **Communications of the ACM**, Volume 38 Issue 11  
Full text available:  [pdf\(400.28 KB\)](#) Additional Information: [full citation](#), [abstract](#), [index terms](#)  
AI as a field has undergone rapid growth in diversification and practicality. For the past 10 years, the repertoire of AI techniques has evolved and expanded. Scores of newer fields have recently been added to the traditional domains of practical AI. Although much practical AI is still best characterized as advanced computing rather than intelligence, applications in everyday commercial and industrial settings have certainly increased, especially since 1990. Additionally, A ...

**30** [Computational methods \("paradigms"\): The typology of unknown words: an experimental study of two corpora](#)  
Xiaobo Ren, François Perrault  
August 1992 **Proceedings of the 14th conference on Computational linguistics - Volume 1**  
Full text available:  [pdf\(464.19 KB\)](#) Additional Information: [full citation](#), [references](#)

**31** [Papers: Extraction of lexical translations from non-aligned corpora](#)  
Kumiko Tanaka, Hideya Iwasaki  
August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 2**  
Full text available:  [pdf\(571.08 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)  
A method for extracting lexical translations from non-aligned corpora is proposed to cope with the unavailability of large aligned corpus. The assumption that "translations of two co-occurring words in a source language also co-occur in the target language" is adopted and represented in the stochastic matrix formulation. The translation matrix provides the co-occurring information translated from the source into the target. This translated co-occurring information should resemble that of the ori ...

**32** [A rule-based hyphenator for Modern Greek](#)  
Theodora I. Noussia  
September 1997 **Computational Linguistics**, Volume 23 Issue 3  
Full text available:  [pdf\(1.16 MB\)](#)  Additional Information: [full citation](#), [abstract](#), [references](#)  
[Publisher Site](#)  
The purpose of this paper is to formally examine hyphenation as it pertains to Modern Greek with the aim of achieving accurate and thorough machine hyphenation. Grammar rules are

interpreted and formally expressed in terms of regular expressions of word substrings, and exact hyphenation rules are derived. Vowel splitting, which traditionally is indicated in terms of prohibitive rather than explicit grammar rules, is examined in detail. Many ambiguities caused by circular definitions of the prohi ...

### **33 CYC, WordNet, and EDR: critiques and responses**

Doug Lenat, George Miller, Toshio Yokoi

November 1995 Communications of the ACM, Volume 38 Issue 11

Full text available:  pdf(106.09 KB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

I applaud Miller's WordNet project and feel that there is much in common in our approaches, even though there are fundamental differences in the two expressions of that spirit. Here, I list the four differences I noted, closing with a crucial observation concerning the common spirit in our work.

34 Association-based natural language processing with neural networks

Kimura Kazuhiro, Suzuoka Takashi, Amano Shin-ya

June 1992 **Proceedings of the 30th annual meeting on Association for Computational Linguistics**

Full text available:  pdf(450.05 KB)

Additional Information: [full citation](#), [abstract](#), [references](#)

 Publisher Site

This paper describes a natural language processing system reinforced by the use of association of words and concepts, implemented as a neural network. Combining an associative network with a conventional system contributes to semantic disambiguation in the process of interpretation. The model is employed within a *kana-kanji conversion system* and the advantages over conventional ones are shown.

## 35 Part-of-speech induction from scratch

Hinrich Schütze

June 1993 **Proceedings of the 31st annual meeting on Association for Computational Linguistics**

Full text available:  pdf(717.90 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Publisher Site

This paper presents a method for inducing the parts of speech of a language and part-of-speech labels for individual words from a large text corpus. Vector representations for the part-of-speech of a word are formed from entries of its near lexical neighbors. A dimensionality reduction creates a space representing the syntactic categories of unambiguous words. A neural net trained on these spatial representations classifies individual contexts of occurrence of ambiguous words. The method classif ...

### 36 Lexicon: Noun phrasal entries in the EDR English word dictionary

A. Koizumi, M. Arioka, C. Harada, M. Sugimoto, L. Guthrie, C. Watts, R. Catizone, Y. Wilks

August 1994 **Proceedings of the 15th conference on Computational linguistics - Volume 1**

Full text available: [pdf\(480.83 KB\)](#) Additional Information: [full citation](#), [references](#)

**Keywords:** lexicon Construction, resources for CL, universal features

37

## Query term disambiguation for Web cross-language information retrieval using a search engine

Akira Maeda, Fatiha Sadat, Masatoshi Yoshikawa, Shunsuke Uemura  
November 2000 **Proceedings of the fifth international workshop on Information  
retrieval with Asian languages**

Full text available:  pdf(736.31 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

With the worldwide growth of the Internet, research on Cross-Language Information Retrieval (CLIR) is being paid much attention. Existing CLIR approaches based on query translation require parallel corpora or comparable corpora for the disambiguation of translated query terms. However, those natural language resources are not readily available. In this paper, we propose a disambiguation method for dictionary-based query translation that is independent of the availability of such scarce langua ...

**Keywords:** WWW, cross-language information retrieval, mutual information, search engines

38 XRel: a path-based approach to storage and retrieval of XML documents using relational databases

August 2001 ACM Transactions on Internet Technology (TOIT), Volume 1 Issue 1

Full text available:  [pdf\(264.27 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

This article describes XRel, a novel approach for storage and retrieval of XML documents using relational databases. In this approach, an XML document is decomposed into nodes on the basis of its tree structure and stored in relational tables according to the node type, with path information from the root to each node. XRel enables us to store XML documents using a fixed relational schema without any information about DTDs and also to utilize indices such as the B+

**Keywords:** XML query, XPath, text markup, text tagging

### **39 Large-scale resources: A Chinese corpus for linguistic research**

Chu-Ren Huang, Keh-jieann Chen

August 1992 Proceedings of the 14th conference on Computational linguistics - Volume 4

Full text available:  pdf(259.12 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This is a project note on the first stage of the construction of a comprehensive corpus of both Modern and Classical Chinese. The corpus is built with the dual aim of serving as the central database for Chinese language processing and for supporting in-depth linguistic research in Mandarin Chinese.

40 Developing hypertext documents for an international audience

Elizabeth S. Spragins

## **November 1992 Proceedings of the 10th annual international conference on Systems documentation**

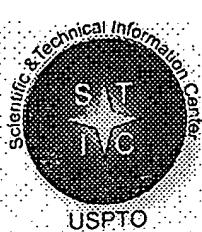
Full text available:  pdf(795.30 KB) Additional Information: [full citation](#), [references](#), [index terms](#)

Results 21 - 40 of 57

Result page: [previous](#) [1](#) [2](#) [3](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright ? 2005 ACM, Inc.  
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

**Useful downloads:**  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

**STIC EIC 2100**

158554

**Search Request Form**Today's Date: July 8, 2005

What date would you like to use to limit the search?

Priority Date: 9/28/2001 Other:Name Leslie Wong

Format for Search Results (Circle One):

AU 2167 Examiner # 18953 PAPER     DISK     EMAILRoom # 3B09 Phone 2-4120 USP     DWPI     EPO     JPO     ACM    IBM TDB  
 IEEE     INSPEC     SPI    Other \_\_\_\_\_Serial # 09/964,453Is this a "Fast & Focused" Search Request? (Circle One)  YES     NOA "Fast & Focused" Search is completed in 2-3 hours (maximum). The search must be on a very specific topic and meet certain criteria. The criteria are posted in EIC2100 and on the EIC2100 NPL Web Page at <http://ptoweb/patents/stic/stic-tc2100.htm>.

What is the topic, novelty, motivation, utility, or other specific details defining the desired focus of this search? Please include the concepts, synonyms, keywords, acronyms, definitions, strategies, and anything else that helps to describe the topic. Please attach a copy of the abstract, background, brief summary, pertinent claims and any citations of relevant art you have found.

Topic: electronic dictionary

Novelty: during a second data request, storing the second context info for the word and displaying the definition of the word and the first context info.

See attached East search history

STIC Searcher Geoffrey St-Leger Phone 83540  
Date picked up 7/8/5 Date Completed 7/8/5

File 347:JAPIO Nov 1976-2005/Feb (Updated 050606)

(c) 2005 JPO & JAPIO

File 350:Derwent WPIX 1963-2005/UD,UM &UP=200543

(c) 2005 Thomson Derwent

Set	Items	Description
S1	22509	DICTIONAR??? OR GLOSSAR???
S2	299805	WORD? ? OR KEYWORD? ? OR TERM? ?
S3	4232	S2(5N) (DEFIN??? OR DEFINITION? ? OR MEANING)
S4	368	S2(7N) CONTEXT???
S5	3605	S2(5N) (ADJACENT OR NEXT()) TO OR NEAR OR NEARBY OR CLOSE OR - AROUND OR SAME() SENTENCE OR PROXIMAL? OR PROXIMITY OR RANGE OR VICINIT??? OR ADJOIN??? OR NEIGHBOR??? OR BESIDE? ?)
S6	2081	(CONTEXT OR S5)(7N) (STOR??? OR SAV??? OR ADD??? OR INSERT?-? ? OR INCORPORAT??? OR INCLUD??? OR SUBMIT???? OR ENTER??? OR INPUT???? OR COPY??? OR COPIE? ?)
S7	141	(S1 OR S3) AND S4:S5 AND S6
S8	117	S7 AND IC=G06F
S9	1462	(ADJACENT OR NEXT()) TO OR NEAR OR NEARBY OR CLOSE OR AROUND OR SAME() SENTENCE OR PROXIMAL? OR PROXIMITY OR (WITHIN OR IN)- () RANGE OR VICINIT??? OR ADJOIN??? OR NEIGHBOR??? OR BESIDE? -? ) (5W) S2
S10	79	(S1 OR S3) AND (S4 OR S9) AND S6
S11	65	S10 AND IC=G06F
S12	14	S10 NOT S11
S13	65	S11
S14	91	SURROUND?(5W) S2
S15	6	S14(7N) (STOR??? OR SAV??? OR ADD??? OR INSERT??? OR INCORP- ORAT??? OR INCLUD??? OR SUBMIT???? OR ENTER??? OR COPY??? OR COPIE? ?)
S16	1	(S1 OR S3) AND S15

12/5/7 (Item 1 from file: 350)  
DIALOG(R) File 350:Derwent WPIX  
(c) 2005 Thomson Derwent. All rts. reserv.

014065490 \*\*Image available\*\*  
WPI Acc No: 2001-549703/200161

XRPX Acc No: N01-408358

Expansion method for effective active vocabulary, involves modifying recognition candidate to substitute proposed word for word fragment and adjacent word fragments used to form proposed word

Patent Assignee: DRAGON SYSTEMS INC (DRAG-N)

Inventor: ALBINA T A; CHEVALIER H L; GILLICK L S; MOORE M B; RENSING P E; YAMRON J P; YOUNG J H

Number of Countries: 095 Number of Patents: 003

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
WO 200109879	A1	20010208	WO 2000US21125	A	20000803	200161 B
AU 200066191	A	20010219	AU 200066191	A	20000803	200161
EP 1204966	A1	20020515	EP 2000953807	A	20000803	200239
			WO 2000US21125	A	20000803	

Priority Applications (No Type Date): US 99390370 A 19990907; US 99147006 P 19990803

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
WO 200109879	A1	E 131	G10L-015/18	

Designated States (National): AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW

Designated States (Regional): AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TZ UG ZW

AU 200066191 A G10L-015/18 Based on patent WO 200109879

EP 1204966 A1 E G10L-015/18 Based on patent WO 200109879

Designated States (Regional): AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI

Abstract (Basic): WO 200109879 A1

NOVELTY - A recognition candidate is modified to substitute a proposed word for the word fragment and the adjacent word fragments or words used to form the proposed word, when the word fragment can be combined with adjacent word fragments to form the proposed word. The proposed word is included in the back up dictionary of the speech recognition system.

DETAILED DESCRIPTION - The recognition candidate is discarded when the word fragment cannot be combined with the adjacent word fragments to form the proposed word included in the back-up dictionary of the speech recognition system. The word fragment is included in the recognition candidate from a speech recognizer (215). The speech recognizer performs speech recognition on a user utterance to produce one or more recognition candidates. INDEPENDENT CLAIMS are also included for the following:

- (a) a method of recognizing speech;
- (b) a method of generating an acoustic model of a word fragment;
- (c) a computer implemented speech recognition system;
- (d) and a computer software.

USE - Used for expanding effective active vocabulary of a speech recognition system.

ADVANTAGE - Enables increasing the effective size of the active vocabulary of the speech recognition system by using fragmented word models. Enables improving the ability of the recognizer to recognize less-frequently used words, since the size of the active vocabulary is increased.

DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of the speech recognition software of the speech recognition system.

Speech recognizer (215)  
pp; 131 DwgNo 2/30

Title Terms: EXPAND; METHOD; EFFECT; ACTIVE; VOCABULARY; MODIFIED;  
RECOGNISE; CANDIDATE; SUBSTITUTE; PROPOSED; WORD; WORD; FRAGMENT;  
ADJACENT; WORD; FRAGMENT; FORM; PROPOSED; WORD

Derwent Class: P86; T01; W04

International Patent Class (Main): G10L-015/18

File Segment: EPI; EngPI

12/5/8 (Item 2 from file: 350)  
DIALOG(R) File 350:Derwent WPIX  
(c) 2005 Thomson Derwent. All rts. reserv.

010907255 \*\*Image available\*\*

WPI Acc No: 1996-404206/199641

XRPX Acc No: N96-340498

Reference pattern training system giving recognition unit smaller than word - executes recognition analysis of subject word, extracts recognition unit, selects unit, reassembles extracted unit in preceeding and succeeding phoneme context , executes collation check and stores unit in accord with phoneme context

Patent Assignee: NEC CORP (NIDE )

Inventor: HATTORI H

Number of Countries: 005 Number of Patents: 005

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
EP 731447	A2	19960911	EP 96103690	A	19960308	199641 B
JP 8248975	A	19960927	JP 9550195	A	19950309	199649
US 5774848	A	19980630	US 96613887	A	19960311	199833
EP 731447	B1	20020612	EP 96103690	A	19960308	200239
DE 69621674	E	20020718	DE 621674	A	19960308	200255
			EP 96103690	A	19960308	

Priority Applications (No Type Date): JP 9550195 A 19950309

Patent Details:

Patent No Kind Lan Pg Main IPC Filing Notes

EP 731447 A2 E 12 G10L-003/00

Designated States (Regional): DE FR GB

JP 8248975 A 11 G10L-003/00

US 5774848 A G10L-005/06

EP 731447 B1 E G10L-015/06

Designated States (Regional): DE FR GB

DE 69621674 E G10L-015/06 Based on patent EP 731447

Abstract (Basic): EP 731447 A

The training system takes out and compares phoneme data about recognition units stored in a recognition word dictionary (10) and a training word dictionary (11). A recognition unit is selected from the recognition training words stored in the training word dictionary . The selected recognition unit reassembles each selected recognition unit in a subject recognition word stored in the recognition word dictionary in phoneme context .

The selected recognition unit is sent to a recognition reference pattern generator (14), which utilises a recognition unit which is best in accord in phoneme context among training data stored in a trainig data memory (12) concerning a recognition unit selected by the training unit selector (13) .

ADVANTAGE - Can expand scope of phoneme context while reducing amount of stored data and provides high recognition performance for speech recognition system.

Dwg.1/5

Title Terms: REFERENCE; PATTERN; TRAINING; SYSTEM; RECOGNISE; UNIT; SMALLER ; WORD; EXECUTE; RECOGNISE; ANALYSE; SUBJECT; WORD; EXTRACT; RECOGNISE; UNIT; SELECT; UNIT; EXTRACT; UNIT; SUCCEEDING; PHONEME; CONTEXT; EXECUTE;

13/5/5 (Item 5 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

06296302 \*\*Image available\*\*  
METHOD AND DEVICE FOR COMPREHENDING LANGUAGE

PUB. NO.: 11-237894 [JP 11237894 A]  
PUBLISHED: August 31, 1999 (19990831)  
INVENTOR(s): NAKANO MIKIO  
APPLICANT(s): NIPPON TELEGR & TELEPH CORP <NTT>  
APPL. NO.: 10-037837 [JP 9837837]  
FILED: February 19, 1998 (19980219)  
INTL CLASS: G10L-003/00; G06F-015/18 ; G06F-017/27

#### ABSTRACT

PROBLEM TO BE SOLVED: To provide a voice interactive system with which a voice having the non-clear partition of speaking can be comprehended in real time and immediate responding is enabled by speaking at free timing.

SOLUTION: When any word is inputted, while referring to a dictionary 250, the semantic description of that word is found by an interactive context managing part 210, a new interactive context candidate is provided together with respective interactive context candidates provided up to that time point, a rule application part 220 obtains a further new interactive context candidate by applying language comprehension rules 110 to these interactive context candidates, a priority calculating part 230 calculates priority by adding these new interactive context candidates to the source interactive context candidates, and the interactive context managing part 210 outputs the interactive context candidate of the highest priority. This processing is repeated each time a word is inputted.

COPYRIGHT: (C)1999,JPO

13/5/6 (Item 6 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

06243281 \*\*Image available\*\*  
TRANSLATION METHOD AND SYSTEM

PUB. NO.: 11-184855 [JP 11184855 A]  
PUBLISHED: July 09, 1999 (19990709)  
INVENTOR(s): NASUKAWA TETSUYA  
TAKEDA KOICHI  
WATANABE HIDEO  
APPLICANT(s): IBM JAPAN LTD  
APPL. NO.: 09-354386 [JP 97354386]  
FILED: December 24, 1997 (19971224)  
INTL CLASS: G06F-017/28

#### ABSTRACT

PROBLEM TO BE SOLVED: To improve the selection accuracy of translated words in a machine translation mode without lowering the processing efficiency by using plural types of dictionaries including a context dictionary when a word that is not defined in a compound word dictionary is translated in a sentence.

SOLUTION: Every input sentence is taken out at its head (110), and the compound words corresponding to a word string composing a single input sentence are retrieved from a compound word dictionary. When each of words which are not corresponding to the compound words is translated (120), its translation is decided by a context dictionary and the

translated word is obtained based on the translation (130). The translated word undergoes the translation result registering processing. Then it's checked whether an object word is stored in a translation result recording buffer as a header. If the object word is stored in the buffer, it's checked whether or not its translated word is stored in the buffer and all words which are not corresponding to the compound words are processed (140, 160). Then all words are translated (170). When it's decided that a full sentence is translated, the sentence is registered (180 to 195) after undergoing the retranslation effect evaluation processing.

COPYRIGHT: (C)1999,JPO

13/5/7 (Item 7 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

05888707 \*\*Image available\*\*  
DEVICE AND METHOD FOR CANCELING SEMANTIC AMBIGUITY

PUB. NO.: 10-171807 [JP 10171807 A]  
PUBLISHED: June 26, 1998 (19980626)  
INVENTOR(s): RI KO  
TAKEUCHI JUNICHI  
APPLICANT(s): NEC CORP [000423] (A Japanese Company or Corporation), JP  
(Japan)  
APPL. NO.: 08-352634 [JP 96352634]  
FILED: December 13, 1996 (19961213)  
INTL CLASS: [6] G06F-017/27 ; G10L-003/00  
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications); 42.5  
(ELECTRONICS -- Equipment)  
JAPIO KEYWORD: R107 (INFORMATION PROCESSING -- OCR & OMR Optical Readers);  
R108 (INFORMATION PROCESSING -- Speech Recognition &  
Synthesis)

#### ABSTRACT

PROBLEM TO BE SOLVED: To attain high performance ambiguity cancelation by deleting a group of an unreliable evidence and word meaning supported by the evidence from a determination list and canceling semantic ambiguity based only on a reliable evidence.

SOLUTION: A determination list learning part 2 calculates logarithmic likelihood ratio between conditional probability values in the appearance of respective word meanings by using information appearing in contexts around a word in question in an input text as evidences and sets up the groups of evidences arranged in the descending order and word meanings supported by the evidences as a determination list. Then mutual information quantity between each evidence and the word in question is calculated and a group of an evidence of which mutual information volume does not exceed a threshold and a word meaning supported by the evidence is deleted from the list and stored in a determination list storing part 3. A semantic ambiguity canceling part 5 successively checks whether each evidence described in the list appears in contexts around the word in question in the input text or not and outputs a word meaning supported by the evidence described in the list as the meaning of the word in question.

13/5/9 (Item 9 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

05808519 \*\*Image available\*\*  
LANGUAGE PROCESSOR AND ITS METHOD

PUB. NO.: 10-091619 [JP 10091619 A]

PUBLISHED: April 10, 1998 (19980410)  
INVENTOR(s): OU ROHIN  
APPLICANT(s): TOSHIBA CORP [000307] (A Japanese Company or Corporation), JP  
(Japan)  
APPL. NO.: 08-243787 [JP 96243787]  
FILED: September 13, 1996 (19960913)  
INTL CLASS: [6] G06F-017/21 ; G06F-017/22  
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications)  
JAPIO KEYWORD: R011 (LIQUID CRYSTALS); R102 (APPLIED ELECTRONICS -- Video  
Disk Recorders, VDR); R131 (INFORMATION PROCESSING --  
Microcomputers & Microprocessors); R139 (INFORMATION  
PROCESSING -- Word Processors)

#### ABSTRACT

PROBLEM TO BE SOLVED: To enable also a user inexperienced in language to input a suitable word during the preparation of a document.  
SOLUTION: Characters inputted through an input part 11 are successively stored in an input character storing area 13b, a word most close to a character string consisting of respective characters stored in the area 13b is retrieved from a word dictionary stored in a dictionary storing area as a neighboring word in each input of a character and the retrieved word is sub-displayed on a sub-display area 14b of a display part 14. After the end of characters equivalent to one word, a word related to the word concerned is retrieved from the word dictionary as a relative word and sub-displayed on the sub-display area 14b of the display part 14. Thereby the contents of the sub-display can be utilized as the help of the character input on the way of character display, and even after the end of a character input for one word, the contents of the sub-display can be utilized as the help of the character input.

13/5/11 (Item 11 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

05279681 \*\*Image available\*\*  
ON-LINE DICTIONARY AND READ UNDERSTANDING SUPPORT SYSTEM UTILIZING SAME

PUB. NO.: 08-235181 [JP 8235181 A]  
PUBLISHED: September 13, 1996 (19960913)  
INVENTOR(s): NIWA YOSHIKI  
NITTA YOSHIHIKO  
APPLICANT(s): HITACHI LTD [000510] (A Japanese Company or Corporation), JP  
(Japan)  
APPL. NO.: 07-039673 [JP 9539673]  
FILED: February 28, 1995 (19950228)  
INTL CLASS: [6] G06F-017/27 ; G06F-017/28  
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications)  
JAPIO KEYWORD: R131 (INFORMATION PROCESSING -- Microcomputers &  
Microprocessors)

#### ABSTRACT

PURPOSE: To speedily know a proper equivalent with small labor by displaying a proper equivalent of the context to a specified word in easy-to-see form for a user at the top or upper part of a synonym display part or by coloring, etc.

CONSTITUTION: When the user specifies a document to be read through an input means 1, a computation program execution means 3 displays the necessary document in a document display area 21 in a display means 2 by using a document display routine 41 and the user beings to read it. If a word whose meaning is unknown is found, the word on the screen is specified with a pen input means 13. The computation program execution means 3 actuates a synonym display routine 42 in response to the specification of the word to open a synonym display area 22 on the screen

nearby the word , and the meaning of the word stored in a dictionary 51 is displayed on the screen. At this time, when the word has plural meanings, the meaning matching the context of the passage is preferentially displayed at an upper part in consideration of the context.

13/5/14 (Item 14 from file: 347)

DIALOG(R) File 347:JAPIO

(c) 2005 JPO & JAPIO. All rts. reserv.

04629813 \*\*Image available\*\*

BILINGUAL DISPLAY METHOD AND DOCUMENT DISPLAY DEVICE AND DIGITAL COPYING DEVICE

PUB. NO.: 06-301713 [JP 6301713 A]

PUBLISHED: October 28, 1994 (19941028)

INVENTOR(s): KAMEDA MASAYUKI  
KATOOKA TAKASHI  
NARITA MASUMI  
OGURO YOSHIHISA  
OONO AYAKO

APPLICANT(s): RICOH CO LTD [000674] (A Japanese Company or Corporation), JP (Japan)

APPL. NO.: 05-111016 [JP 93111016]

FILED: April 13, 1993 (19930413)

INTL CLASS: [5] G06F-015/38

JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications); 30.2 (MISCELLANEOUS GOODS -- Sports & Recreation)

JOURNAL: Section: , Section No. FFFFFFF, Vol. 94, No. 10, Pg. FFFFFF, FF, FFFF (FFFFFF)

#### ABSTRACT

PURPOSE: To clearly display the correspondence of a translated word of second language corresponding to a word (including an idiom) of first language, in the bilingual display for displaying the translated word under the word of a sentence of first language.

CONSTITUTION: A bilingual word dictionary 1 gives a translated word of second language to a word of first language. A translated word acquiring means 2 retrieves the bilingual word dictionary based on a character-string in a document of first language, and obtains a translated word of second language to the word of first language. A bilingual display means 3 adds and displays the translated word between lines in the vicinity of its position at every word whose translated word is acquired, together with the document of first language. With such a constitution, a range of the word to which the translated word is added is discriminated and displayed.

13/5/19 (Item 19 from file: 347)

DIALOG(R) File 347:JAPIO

(c) 2005 JPO & JAPIO. All rts. reserv.

03848064 \*\*Image available\*\*

DICTIONARY CONSULTING SYSTEM

PUB. NO.: 04-213164 [JP 4213164 A]

PUBLISHED: August 04, 1992 (19920804)

INVENTOR(s): YANO TAKANORI  
MATSUBARA AKIO  
KATOOKA TAKASHI

APPLICANT(s): RICOH CO LTD [000674] (A Japanese Company or Corporation), JP (Japan)

APPL. NO.: 03-044160 [JP 9144160]

FILED: February 15, 1991 (19910215)

INTL CLASS: [5] G06F-015/38 ; G06F-015/20  
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications); 30.2  
(MISCELLANEOUS GOODS -- Sports & Recreation)  
JOURNAL: Section: P, Section No. 1455, Vol. 16, No. 559, Pg. 83,  
November 30, 1992 (19921130)

ABSTRACT

PURPOSE: To appropriately translate a word by translating the word from the context of the designated word, and checking the part of a text.

CONSTITUTION: When there are an unknown word during a text input (edition), and the pertinent unknown word is desired to be translated, the word is indicated by an inputting device 10, and a dictionary consultation processor 8 is activated. Then, an analysis is executed for the range of the text including the word, and the appropriate meaning of the pertinent word is displayed. And also, a translated word is displayed after or under the designated word.

13/5/20 (Item 20 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

03808969 \*\*Image available\*\*  
LANGUAGE UNDERSTANDING ASSISTING DEVICE

PUB. NO.: 04-174069 [JP 4174069 A]  
PUBLISHED: June 22, 1992 (19920622)  
INVENTOR(s): NARITA MASUMI  
APPLICANT(s): RICOH CO LTD [000674] (A Japanese Company or Corporation), JP  
(Japan)  
APPL. NO.: 02-293002 [JP 90293002]  
FILED: October 30, 1990 (19901030)  
INTL CLASS: [5] G06F-015/38 ; G06F-003/14  
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications); 30.2  
(MISCELLANEOUS GOODS -- Sports & Recreation); 45.3  
(INFORMATION PROCESSING -- Input Output Units)  
JAPIO KEYWORD: R107 (INFORMATION PROCESSING -- OCR & OMR Optical Readers)  
JOURNAL: Section: P, Section No. 1433, Vol. 16, No. 489, Pg. 4,  
October 09, 1992 (19921009)

ABSTRACT

PURPOSE: To speedily understand a document by consulting with a dictionary immediately when a word whose meaning can not be understood is encountered at the time of understanding the document, and displaying dictionary data nearby the word and storing the data relatively.

CONSTITUTION: The document in object language is inputted through an input device 10 and the inputted document and the processing result of a processor 1 are displayed on a display device 6. The words in the document displayed on the display device 6 are indicated by a word indication device and looked up in the dictionary by a dictionary retrieval device 8. Dictionary information on the retrieved words is displayed on the display device 6 and also displayed nearby the words by a display controller 5. The words and desired information are related and stored in a memory 2. Consequently, the document can speedily be understood.

13/5/24 (Item 24 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

03435265 \*\*Image available\*\*  
POLYSEMY ELIMINATING DEVICE

PUB. NO.: 03-098165 [JP 3098165 A]  
PUBLISHED: April 23, 1991 (19910423)  
INVENTOR(s): TANI KENICHI  
APPLICANT(s): NEC CORP [000423] (A Japanese Company or Corporation), JP  
(Japan)  
APPL. NO.: 01-236554 [JP 89236554]  
FILED: September 11, 1989 (19890911)  
INTL CLASS: [5] G06F-015/38  
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications); 30.2  
(MISCELLANEOUS GOODS -- Sports & Recreation)  
JOURNAL: Section: P, Section No. 1229, Vol. 15, No. 287, Pg. 56, July  
22, 1991 (19910722)

#### ABSTRACT

PURPOSE: To obtain a powerful polysemy eliminating function at an early stage of analysis by using the limit conditions of adjacency, the stage information, and the priority information which are stored in a dictionary

CONSTITUTION: A set between the conditions and the intensity of the attributes of the meaning of words which are difficult to be placed adjacent to each other in a sentence and a set of the conditions and the intensity of the attributes of the meaning of words adjacent to each other in a sentence are registered into a dictionary 102. At the same time, the priority information is also registered in the dictionary 102 together with the stage information which decides a stage where the polysemy can be eliminated. An input sentence dividing part 101 divides an input sentence into sections of adjacent words by reference to the dictionary 102. An adjacent condition application part 103 deletes the sets satisfying the conditions of the meaning of words which are difficult to be placed adjacent to each other to plural candidate words of the adjacent word sections and the sets which do not satisfy the conditions of the words adjacent to each other in the order of stronger conditions. Then a priority application part 104 selects the meaning of words having high priority to the polysemy that cannot be eliminated at the subsequent stages based on the stage information. As a result, it is possible to obtain a powerful polysemy eliminating function in an early stage of analysis.

13/5/29 (Item 29 from file: 347)  
DIALOG(R) File 347:JAPIO  
(c) 2005 JPO & JAPIO. All rts. reserv.

02454874 \*\*Image available\*\*  
CONTEXT ANALYZING SYSTEM

PUB. NO.: 63-071774 [JP 63071774 A]  
PUBLISHED: April 01, 1988 (19880401)  
INVENTOR(s): KAMEI SHINICHIRO  
APPLICANT(s): NEC CORP [000423] (A Japanese Company or Corporation), JP  
(Japan)  
APPL. NO.: 61-216510 [JP 86216510]  
FILED: September 12, 1986 (19860912)  
INTL CLASS: [4] G06F-015/38  
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications); 30.2  
(MISCELLANEOUS GOODS -- Sports & Recreation)  
JOURNAL: Section: P, Section No. 744, Vol. 12, No. 298, Pg. 142,  
August 15, 1988 (19880815)

#### ABSTRACT

PURPOSE: To presume the logical structure and sentence structure of a composition along the intention of the writer of the composition before making the syntax analysis and semantic analysis, by providing a context form holding mechanism, context structure collating mechanism, and context

structure presuming mechanism.

CONSTITUTION: An inputted composition read by an input composition reading section 1 is transmitted to a form element analyzing section 2 where the composition is divided into words on the basis of a **dictionary** 3, and a content of the **dictionary** 3 are assigned to each word and transmitted to a **context** analyzing section 4. The context analyzing section 4 is composed of a context form holding mechanism 5, context structure collating mechanism 6, and context structure presuming mechanism 7 and the context structure collating mechanism 6 detects candidates of the **context** structure of the inputted composition by referring to the information of vocabularies indicating a **context stored** in the **context** form holding mechanism 5. The context structure presuming mechanism 7 presumes the most probable context structure out of the received context structure candidates and a syntax analyzing section 8 carries out a syntax analysis in accordance with the presumed context structure. A semantic analyzing section 9 extracts the meaning of the inputted composition on the basis of the structure of the inputted composition obtained as a result of the syntax analysis.

13/5/33 (Item 2 from file: 350)  
DIALOG(R) File 350:Derwent WPIX  
(c) 2005 Thomson Derwent. All rts. reserv.

016597709 \*\*Image available\*\*  
WPI Acc No: 2004-756443/200474  
XRPX Acc No: N04-597345

Handheld electronic book reader annotating method for use in educational environment, involves creating context for defined term in electronic document by associating selected portion of electronic document with defined term  
Patent Assignee: INT BUSINESS MACHINES CORP (IBMC )  
Inventor: BARSNESS E L; SANTOSUSSO J M  
Number of Countries: 001 Number of Patents: 001  
Patent Family:  
Patent No Kind Date Applicat No Kind Date Week  
US 20040201633 A1 20041014 US 2001951363 A 20010913 200474 B

Priority Applications (No Type Date): US 2001951363 A 20010913

Patent Details:

Patent No Kind Lan Pg Main IPC Filing Notes  
US 20040201633 A1 14 G09G-005/00

Abstract (Basic): US 20040201633 A1

NOVELTY - The method involves selecting a portion of an electronic document displayed in a handheld electronic book reader in response to user input . A **context** is created for a **defined term** in the electronic document by associating the selected portion of the electronic document with the **defined term** . The definition of the **defined term** and the **context** associated with the **term** is displayed in response to user input on a display (14).

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(A) a method of displaying an electronic document on a handheld electronic book reader

(B) a method of monitoring usage of an electronic document.

USE - Used for annotating with a handheld electronic book reader that is utilized in educational or classroom environment by student and instructor, and for general entertainment or informational purpose.

ADVANTAGE - The method enhances the collaborative and annotative capabilities of electronic book readers and hence provide useful feedback on reader using the electronic documents e.g. to improve lesson plans associated with such documents, to monitor student activity, and to identify students needing supplemental assistance. The

method provides significant benefits in terms of improving the educational process for both students and instructors.

DESCRIPTION OF DRAWING(S) - The drawing shows a top plan view of a handheld electronic book reader incorporating annotation and usage tracking capabilities using the handheld electronic book reader annotating method.

Reader (12)  
Touch screen display (14)  
Power button (15)  
Menu button (16)  
Electronic book reader (20)

pp; 14 DwgNo 1/7

Title Terms: ELECTRONIC; BOOK; READ; METHOD; EDUCATION; ENVIRONMENT; CONTEXT; DEFINE; TERM; ELECTRONIC; DOCUMENT; ASSOCIATE; SELECT; PORTION; ELECTRONIC; DOCUMENT; DEFINE; TERM

Derwent Class: P85; T01; W04

International Patent Class (Main): G09G-005/00

International Patent Class (Additional): G06F-017/24

File Segment: EPI; EngPI

13/5/36 (Item 5 from file: 350)  
DIALOG(R) File 350:Derwent WPIX  
(c) 2005 Thomson Derwent. All rts. reserv.

014868696 \*\*Image available\*\*  
WPI Acc No: 2002-689402/200274

XRPX Acc No: N02-543694

Term context learning method in terminological system, involves selecting focal point category that identifies concept for category relationship for term to learn from base term, based on assigned weight values

Patent Assignee: ORACK CORP (ORAC-N)

Inventor: CONKLIN J

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
US 6415283	B1	20020702	US 98170895	A	19981013	200274 B

Priority Applications (No Type Date): US 98170895 A 19981013

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 6415283	B1	20		G06F-017/30	

Abstract (Basic): US 6415283 B1

NOVELTY - The selected category nodes representing concepts for category relationships, are assigned with weight values associated with corresponding base terms. A cluster of categories of nodes is selected based on assigned values and category relationships. A focal point category identifying a concept of a term to learn from a base term based on assigned values, is selected for the cluster.

DETAILED DESCRIPTION - An INDEPENDENT CLAIM is included for computer readable medium storing term context learning program.

USE - Used to learn the context or meaning of terms by identifying categories from a knowledge catalog in terminological learning system. Used to learn the context of the terms, also for use in control theory, medical diagnostics, criminal profile and fraud detection.

ADVANTAGE - Terms are mapped to categories of a classification system and the clustering techniques are used to identify categories in the system that best reflect the terms input to the terminological system.

DESCRIPTION OF DRAWING(S) - The figure shows the block diagram illustrating the cluster processing.

pp; 20 DwgNo 1/9

Title Terms: TERM; CONTEXT; LEARNING; METHOD; SYSTEM; SELECT; FOCUS; POINT;  
CATEGORY; IDENTIFY; CONCEPT; CATEGORY; RELATED; TERM; LEARNING; BASE;  
TERM; BASED; ASSIGN; WEIGHT; VALUE  
Derwent Class: T01  
International Patent Class (Main): G06F-017/30  
File Segment: EPI

13/5/37 (Item 6 from file: 350)  
DIALOG(R) File 350:Derwent WPIX  
(c) 2005 Thomson Derwent. All rts. reserv.

014552010  
WPI Acc No: 2002-372713/200241  
XRPX Acc No: N02-291240  
**Complicated context correlation processing technique**  
Patent Assignee: HUAJIAN MACHINE TEANSLATION CO LTD (HUAJ-N); CHEN Z  
(CHEN-I)  
Inventor: CHEN Z  
Number of Countries: 001 Number of Patents: 002  
Patent Family:  

Patent No	Kind	Date	Applicat No	Kind	Date	Week
CN 1180203	A	19980429	CN 97111944	A	19970702	200241 B
CN 1067781	C	20010627	CN 97111944	A	19970702	200504

Priority Applications (No Type Date): CN 97111944 A 19970702

Patent Details:  
Patent No Kind Lan Pg Main IPC Filing Notes  
CN 1180203 A G06F-017/28  
CN 1067781 C G06F-017/28

Abstract (Basic): CN 1180203 A

NOVELTY - The present invention includes the following technical steps: 1. embedding **context** -dependent information and context-dependent operation in rule and **dictionary**, the form of rule is header--context-dependent function, right portion and conversion body; the form of every word in the **dictionary** is: entry **word**, characteristic set, **context** -dependent function and version; 2. for every rule, making header matching first, if the matching is successful, executing context-dependent function of rule to determine context-dependent condition of current header pattern, if said condition is tenable, inducing the content in the current pattern; and 3. in a similar way to rule processing mode, making **context** -dependent processing of **words**. By adopting data and operation integrated technology, it effectively solves the complex context-dependent processing problem.

DwgNo 0/0  
Title Terms: COMPLICATED; CONTEXT; CORRELATE; PROCESS; TECHNIQUE  
Derwent Class: T01  
International Patent Class (Main): G06F-017/28  
File Segment: EPI

13/5/40 (Item 9 from file: 350)  
DIALOG(R) File 350:Derwent WPIX  
(c) 2005 Thomson Derwent. All rts. reserv.

012769718 \*\*Image available\*\*  
WPI Acc No: 1999-575941/199949  
XRPX Acc No: N99-425100  
**Keyword extracting method for use in document search processing - involves semantically classifying corresponding keyword extracted, based on context of speech and meaning of specific word in input sentence**  
Patent Assignee: NIPPON TELEGRAPH & TELEPHONE CORP (NITE )

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 11250097	A	19990917	JP 9853889	A	19980305	199949 B

Priority Applications (No Type Date): JP 9853889 A 19980305

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 11250097	A	9	G06F-017/30		

Abstract (Basic): JP 11250097 A

NOVELTY - Importance degree of a specific word in an input sentence is judged. Based on the context of the speech and the meaning of the word, a corresponding keyword is extracted and classified, semantically. Based on the importance degree of the classified keyword, the keyword of a preset number is output. DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for keyword extracting apparatus.

USE - For use in document search processing.

ADVANTAGE - Since importance degree of keyword is decided, by frequency of word, search accuracy is raised.

Dwg.1/12

Title Terms: KEYWORD; EXTRACT; METHOD; DOCUMENT; SEARCH; PROCESS; CLASSIFY; CORRESPOND; KEYWORD; EXTRACT; BASED; CONTEXT; SPEECH; MEANING; SPECIFIC; WORD; INPUT; SENTENCE

Derwent Class: T01

International Patent Class (Main): G06F-017/30

File Segment: EPI

13/5/41 (Item 10 from file: 350)

DIALOG(R) File 350:Derwent WPIX

(c) 2005 Thomson Derwent. All rts. reserv.

012769683 \*\*Image available\*\*

WPI Acc No: 1999-575906/199949

XRPX Acc No: N99-425065

Input information based context processing apparatus for language processor, machine translator, word processor - has matching relation judging unit which detects whether matching data satisfy predefined conditions and stores them, in context information memory unit

Patent Assignee: MATSUSHITA DENKI SANGYO KK (MATU )

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 11250060	A	19990917	JP 9850456	A	19980303	199949 B

Priority Applications (No Type Date): JP 9850456 A 19980303

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 11250060	A	11	G06F-017/27		

Abstract (Basic): JP 11250060 A

NOVELTY - A search unit (5) searches for matching data in the dictionary (4) and stores it in memory (6). An updating unit (7) updates the stored context data based on an indication from the controller. A matching relation judging unit detects whether the matching data satisfies specific conditions, based on which it is stored in context information memory (9). DETAILED DESCRIPTION - A sentence is input and stored in a memory (2). A controller (3) outputs context information indicating sequentially read words of input sentence, based on which a dictionary (4) stores matching data. The matching data has information about relationship between the group of words that appear together in sentence.

USE - In language processor, machine translator, word

processor, interactive processor, etc.

ADVANTAGE - Accuracy of context information is improved and exact information is output. DESCRIPTION OF DRAWING(S) - The figure shows block diagram of context processing apparatus. (2) Memory; (3) Controller; (4) Dictionary ; (5) Search unit; (6) Memory; (7) Updating unit; (9) Context information memory.

Dwg.1/11

Title Terms: INPUT; INFORMATION; BASED; CONTEXT; PROCESS; APPARATUS; LANGUAGE; PROCESSOR; MACHINE; TRANSLATION; WORD; PROCESSOR; MATCH; RELATED; JUDGEMENT; UNIT; DETECT; MATCH; DATA; SATISFY; PREDEFINED; CONDITION; STORAGE; CONTEXT; INFORMATION; MEMORY; UNIT

Derwent Class: T01

International Patent Class (Main): G06F-017/27

International Patent Class (Additional): G06F-017/28

File Segment: EPI

13/5/42 (Item 11 from file: 350)  
DIALOG(R) File 350:Derwent WPIX  
(c) 2005 Thomson Derwent. All rts. reserv.

012697844 \*\*Image available\*\*

WPI Acc No: 1999-503953/199942

XRPX Acc No: N99-376764

Context processing apparatus in machine translation apparatus, interactive processing apparatus, word processor - includes updating processing unit which updates context information stored in respective memory only when words in context information memory corresponds to stored coincidence data

Patent Assignee: MATSUSHITA DENKI SANGYO KK (MATU )

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 11219360	A	19990810	JP 9820572	A	19980202	199942 B

Priority Applications (No Type Date): JP 9820572 A 19980202

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 11219360	A	9		G06F-017/27	

Abstract (Basic): JP 11219360 A

NOVELTY - The updating processing unit (7) updates context information stored in the memory (9) only when the words in the context information memory corresponds to coincidence data stored in memory (6) after reading all the words in input sentence. DETAILED DESCRIPTION - The memory (2) stores input sentence. The dictionary (4) stores coincidence data. The searching unit searches coincidence data from the coincidence dictionary corresponding to search key. The memory (6) stores the searched coincidence data. The updating processing unit updates the context information. The judging unit (8) judges if the group of arbitrary words is stored in the coincidence dictionary as coincidence data. The context information memory stores context information along with corresponding coincidence data.

USE - For context processing in machine translation apparatus, interactive processing apparatus, word processor.

ADVANTAGE - Facilitates to process context information with high efficiency by confirming predefined relationship between coincidence data memory and context data memory. DESCRIPTION OF DRAWING(S) - The figure depicts functional block diagram of context processing apparatus. (2,6,9) Memories; (4) Dictionary ; (7) Updating processing unit; (8) Judging unit.

File 8:Ei Compendex(R) 1970-2005/Jun W4  
     (c) 2005 Elsevier Eng. Info. Inc.  
 File 35:Dissertation Abs Online 1861-2005/Jun  
     (c) 2005 ProQuest Info&Learning  
 File 65:Inside Conferences 1993-2005/Jul W1  
     (c) 2005 BLDSC all rts. reserv.  
 File 2:INSPEC 1969-2005/Jun W4  
     (c) 2005 Institution of Electrical Engineers  
 File 94:JICST-EPlus 1985-2005/May W3  
     (c) 2005 Japan Science and Tech Corp(JST)  
 File 6:NTIS 1964-2005/Jun W4  
     (c) 2005 NTIS, Intl Cpyrgh All Rights Res  
 File 144:Pascal 1973-2005/Jun W4  
     (c) 2005 INIST/CNRS  
 File 434:SciSearch(R) Cited Ref Sci 1974-1989/Dec  
     (c) 1998 Inst for Sci Info  
 File 34:SciSearch(R) Cited Ref Sci 1990-2005/Jul W1  
     (c) 2005 Inst for Sci Info  
 File 99:Wilson Appl. Sci & Tech Abs 1983-2005/May  
     (c) 2005 The HW Wilson Co.  
 File 266:FEDRIP 2005/Jun  
     Comp & dist by NTIS, Intl Copyright All Rights Res  
 File 95:TEME-Technology & Management 1989-2005/May W5  
     (c) 2005 FIZ TECHNIK  
 File 583:Gale Group Globalbase(TM) 1986-2002/Dec 13  
     (c) 2002 The Gale Group  
 File 438:Library Lit. & Info. Science 1984-2005/May  
     (c) 2005 The HW Wilson Co  
 File 256:TecInfoSource 82-2005/May  
     (c) 2005 Info.Sources Inc

Set	Items	Description
S1	39303	DICTIONAR??? OR GLOSSAR???
S2	3124983	WORD? ? OR KEYWORD? ? OR TERM? ?
S3	57906	S2(5N) (DEFIN??? OR DEFINITION? ? OR MEANING)
S4	9446	S2(5N) CONTEXT???
S5	26720	(ADJACENT OR NEXT OR NEAR OR NEARBY OR CLOSE OR AROUND OR - SAME()SENTENCE OR PROXIMAL? OR PROXIMITY OR (WITHIN OR IN)()R- ANGE OR VICINIT??? OR ADJOIN??? OR NEIGHBOR??? OR BESIDE? ? OR SURROUND???) (5W) S2
S6	14666	(CONTEXT OR S5)(7N) (STOR??? OR SAV??? OR ADD??? OR INSERT?- ? ? OR INCORPORAT??? OR INCLUD??? OR SUBMIT???? OR ENTER??? OR COPY??? OR COPIE? ?)
S7	104	(S1 OR S3) AND S4:S5 AND S6
S8	79	RD (unique items)
S9	71	S8 NOT PY=2002:2005
S10	69	S9 NOT PD=2002:2005

10/5/3 (Item 3 from file: 8)  
DIALOG(R) File 8:EI Compendex(R)  
(c) 2005 Elsevier Eng. Info. Inc. All rts. reserv.

05219139 E.I. No: EIP99024553721  
Title: Efficient dictionary access method for morphological analysis  
Author: Ando, Kazuaki; Tsuji, Takako; Fuketa, Masao; Aoe, Jun-ichi  
Corporate Source: Univ of Tokushima, Tokushima-Shi, Jpn  
Conference Title: Proceedings of the 1998 IEEE International Conference  
on Systems, Man, and Cybernetics. Part 3 (of 5)  
Conference Location: San Diego, CA, USA Conference Date:  
19981011-19981014  
Sponsor: IEEE  
E.I. Conference No.: 49610  
Source: Proceedings of the IEEE International Conference on Systems, Man  
and Cybernetics 3 1998. IEEE, Piscataway, NJ, USA, 98CB36218. p 2876-2881  
Publication Year: 1998  
CODEN: PICYE3 ISSN: 1062-922X  
Language: English  
Document Type: CA; (Conference Article) Treatment: G; (General Review)  
Journal Announcement: 9903WS  
Abstract: This paper proposes an efficient dictionary access method for  
morphological analysis of oriental languages by extending an Aho and  
Corasick's pattern matching machine. The proposed method is a simple and  
efficient algorithm to find all possible substrings in an input sentence  
and during a single pass. It stores the relations of grammatical  
connectivity of adjacent words into the output functions. Therefore,  
the costs of checking connections between the adjacent words can be  
reduced by using the connectivity relations. Furthermore, the construction  
method of the relations of grammatical connectivity is described. Finally,  
the proposed method is verified by theoretical analysis and an experimental  
estimation is supported by the computer simulation with a 100,000 words  
dictionary. From the simulation results, it turns out that the proposed  
method was 49.9% faster (CPU time) than the traditional trie approach. As  
for the number of candidates for checking connections, it was 25.5% less  
than that of the original morphological analysis. (Author abstract) 11  
Refs.  
Descriptors: \*Natural language processing systems; Pattern matching;  
Pattern recognition systems; Algorithms; Functions; Computer simulation;  
Mathematical morphology  
Identifiers: Dictionary access method  
Classification Codes:  
723.2 (Data Processing); 723.5 (Computer Applications)  
723 (Computer Software); 921 (Applied Mathematics)  
72 (COMPUTERS & DATA PROCESSING); 92 (ENGINEERING MATHEMATICS)

10/5/39 (Item 6 from file: 2)  
DIALOG(R) File 2:INSPEC  
(c) 2005 Institution of Electrical Engineers. All rts. reserv.

5291885 INSPEC Abstract Number: C9607-1230-089  
Title: Syntax, semantics, and pragmatics of contexts  
Author(s): Sowa, J.F.  
Author Affiliation: Philos. & Comput. & Cognitive Sci., State Univ. of  
New York, Binghamton, NY, USA  
Conference Title: Formalizing Context. Papers from the 1995 AAAI Fall  
Symposium. (Tech. Report FS-95-02) p.85-96  
Publisher: AAAI Press, Menlo Park, CA, USA  
Publication Date: 1995 Country of Publication: USA v+119 pp.  
ISBN: 0 929280 96 2 Material Identity Number: XX96-01552  
Conference Title: Formalizing Context. Papers from the 1995 AAAI Fall  
Symposium. (Tech. Report FS-95-02)  
Conference Date: 10-12 Nov. 1995 Conference Location: Cambridge, MA,  
USA

Language: English Document Type: Conference Paper (PA)

Treatment: Theoretical (T)

Abstract: The notion of context is indispensable in discussions of meaning, but the word context has often been used in conflicting senses. In logic, the first representation of context as a formal object was by the philosopher C.S. Peirce; but for nearly eighty years, his treatment was unknown outside a small group of Peirce aficionados. In the early 1980s, three new theories included related notions of context: Kamp's (1981) discourse representation theory; Barwise and Perry's (1983) situation semantics; and Sowa's (1984) conceptual graphs, which explicitly introduced Peirce's approach to the AI community. More recently, John McCarthy and his students have begun to use a closely related notion of context as a basis for organizing and partitioning knowledge bases. Each of the theories has distinctive, but complementary ideas that can enrich the others, but the relationships between them are far from clear. This paper analyzes the semantic foundations of these theories and shows how McCarthy's *ist(c, p)* predicate can be interpreted in terms of the semantic notions underlying the others. (24 Refs)

Subfile: C

Descriptors: artificial intelligence; computational linguistics; formal logic

Identifiers: syntax; semantics; contexts; formal logic; meaning; AI; discourse representation; situation semantics; conceptual graphs; artificial intelligence; knowledge bases; semantic foundations; semantic notions

Class Codes: C1230 (Artificial intelligence); C4210L (Formal languages and computational linguistics); C6170 (Expert systems)

Copyright 1996, IEE

10/5/41 (Item 8 from file: 2)  
DIALOG(R) File 2:INSPEC  
(c) 2005 Institution of Electrical Engineers. All rts. reserv.

5007804 INSPEC Abstract Number: C9509-7820-010

Title: The COFREL project: compiling an Old French-English dictionary

Author(s): Langley, F.

Author Affiliation: Hull Univ., UK

Journal: Literary & Linguistic Computing vol.10, no.1 p.45-9

Publication Date: 1995 Country of Publication: UK

CODEN: LLCOEI ISSN: 0268-1145

Language: English Document Type: Journal Paper (JP)

Treatment: Practical (P)

Abstract: Describes the origins, nature and development of the COFREL (Computerized Old French-English Lexicon) project, whose purpose is the production of an Old French-English dictionary. In part, it is the story of how a team of near-illiterates (in computing terms) conceived of a project in ignorance of both the true potentials of information technology in lexicology and lexicography, and of the dangers of that very ignorance. The paper summarizes the software and hardware employed as the project has progressed, as well as some of the computing techniques used in the collection, storage and preparation of data. It aims to illustrate how, in a long-term project, time and effort can sometimes be misdirected, and how a research strategy should remain flexible enough to take into account technological progress and increasing technical competence on the part of the researchers. (0 Refs)

Subfile: C

Descriptors: glossaries ; language translation

Identifiers: COFREL project; French-English dictionary ; Computerized lexicon; near-illiterates; information technology; lexicology; lexicography ; flexible research strategy; technological progress; technical competence

Class Codes: C7820 (Humanities computing)

Copyright 1995, IEE

10/5/44 (Item 11 from file: 2)  
DIALOG(R)File 2:INSPEC  
(c) 2005 Institution of Electrical Engineers. All rts. reserv.

04180085 INSPEC Abstract Number: C9208-6180N-007  
**Title:** Using multiple knowledge sources for word sense discrimination

Author(s): McRoy, S.W.

Author Affiliation: GE Res. & Dev. Center, Schenectady, NY, USA

Journal: Computational Linguistics vol.18, no.1 p.1-30

Publication Date: March 1992 Country of Publication: USA

CODEN: CLINEE ISSN: 0891-2017

Language: English Document Type: Journal Paper (JP)

Treatment: Practical (P)

**Abstract:** The paper addresses the problem of how to identify the intended meaning of individual words in unrestricted texts, without necessarily having access to complete representations of sentences. To discriminate senses, an understander can consider a diversity of information, including syntactic tags, word frequencies, collocations, semantic context, role-related expectations, and syntactic restrictions. However, current approaches make use of only small subsets of this information. The author describes how to use the whole range of information. The discussion includes how the preference cues relate to general lexical and conceptual knowledge and to more specialized knowledge of collocations and contexts. She describes a method of combining cues on the basis of their individual specificity, rather than a fixed ranking among cue-types. She also discusses an application of the approach in a system that computes sense tags for arbitrary texts, even when it is unable to determine a single syntactic or semantic representation for some sentences. (45 Refs)

Subfile: C

Descriptors: computational linguistics; knowledge representation; natural languages

Identifiers: NLP; multiple knowledge sources; word sense discrimination; intended meaning; unrestricted texts; syntactic tags; word frequencies; semantic context; role-related expectations; syntactic restrictions; preference cues; conceptual knowledge; cues; specificity; sense tags; arbitrary texts

Class Codes: C6180N (Natural language processing); C6170 (Expert systems )

10/5/47 (Item 14 from file: 2)  
DIALOG(R)File 2:INSPEC  
(c) 2005 Institution of Electrical Engineers. All rts. reserv.

00987162 INSPEC Abstract Number: C76030546

**Title:** The data dictionary /directory (DD/D)

Author(s): Delport, L.

Journal: Informatie vol.18, no.7-8 p.446-55

Publication Date: July-Aug. 1976 Country of Publication: Netherlands

CODEN: INFTRC ISSN: 0019-9907

Language: Dutch Document Type: Journal Paper (JP)

Treatment: Practical (P)

**Abstract:** Data dictionaries /directories (DD/D) are defined and described. The main features include the complication of references with unambiguous listings and facilities for cross-indexing. Potential users are listed along with the associated activities. The structure of a typical DD/D is a hierarchical system with segments for catalogue components. The KWIC index (key-word-in-context) is applicable and COPY routines are available. About six examples are given of commercially available data dictionaries along the lines discussed. A useful feature is the inclusion of a glossary to elaborate special terms used in the text, e.g. HIDAM=hierarchical indexed direct access method.

File 275:Gale Group Computer DB(TM) 1983-2005/Jul 07  
     (c) 2005 The Gale Group  
 File 621:Gale Group New Prod.Annou.(R) 1985-2005/Jul 08  
     (c) 2005 The Gale Group  
 File 636:Gale Group Newsletter DB(TM) 1987-2005/Jul 07  
     (c) 2005 The Gale Group  
 File 16:Gale Group PROMT(R) 1990-2005/Jul 07  
     (c) 2005 The Gale Group  
 File 160:Gale Group PROMT(R) 1972-1989  
     (c) 1999 The Gale Group  
 File 148:Gale Group Trade & Industry DB 1976-2005/Jul 08  
     (c) 2005 The Gale Group  
 File 624:McGraw-Hill Publications 1985-2005/Jul 08  
     (c) 2005 McGraw-Hill Co. Inc  
 File 15:ABI/Inform(R) 1971-2005/Jul 08  
     (c) 2005 ProQuest Info&Learning  
 File 647:CMP Computer Fulltext 1988-2005/Jun W3  
     (c) 2005 CMP Media, LLC  
 File 674:Computer News Fulltext 1989-2005/Jul W1  
     (c) 2005 IDG Communications  
 File 696:DIALOG Telecom. Newsletters 1995-2005/Jun 20  
     (c) 2005 The Dialog Corp.  
 File 369:New Scientist 1994-2005/May W2  
     (c) 2005 Reed Business Information Ltd.  
 File 810:Business Wire 1986-1999/Feb 28  
     (c) 1999 Business Wire  
 File 813:PR Newswire 1987-1999/Apr 30  
     (c) 1999 PR Newswire Association Inc  
 File 610:Business Wire 1999-2005/Jul 08  
     (c) 2005 Business Wire.  
 File 613:PR Newswire 1999-2005/Jul 08  
     (c) 2005 PR Newswire Association Inc

Set	Items	Description
S1	114516	DICTIONAR??? OR GLOSSAR???
S2	8423908	WORD? ? OR KEYWORD? ? OR TERM? ?
S3	226206	S2 (5N) (DEFIN??? OR DEFINITION? ? OR MEANING)
S4	18051	S2 (5N) CONTEXT???
S5	320063	(ADJACENT OR NEXT OR NEAR OR NEARBY OR CLOSE OR AROUND OR - SAME()SENTENCE OR PROXIMAL? OR PROXIMITY OR (WITHIN OR IN) ()R- ANGE OR VICINIT??? OR ADJOIN??? OR NEIGHBOR??? OR BESIDE? ? OR SURROUND???) (5W) S2
S6	50893	(CONTEXT OR S5) (7N) (STOR??? OR SAV??? OR ADD??? OR INSERT?- ?? OR INCORPORAT??? OR INCLUD??? OR SUBMIT???? OR ENTER??? OR COPY??? OR COPIE? ?)
S7	513	(S1 OR S3) (50N) S4:S5 (50N) S6
S8	10447	S1 (5N) (BUILD??? OR BUILT OR CONSTRUCT??? OR CREAT??? OR GE- NERAT??? OR PRODUC??? OR ESTABLISH?)
S9	2542	(PERSONAL OR CUSTOM? OR ALTERNATE OR ALTERNATIVE) (3W) S1
S10	28	(S1 OR S3) (50N) S4:S5 (50N) S6 (50N) S8:S9
S11	22	RD (unique items)
S12	1106	USER(2W) S1
S13	12	(S1 OR S3) (50N) S4:S5 (50N) S6 (50N) S12
S14	6	RD (unique items)
S15	102	S6 (7N) S1
S16	52	S15 (50N) S4:S5
S17	38	RD (unique items)
S18	23	S17 NOT (S11 OR S14 OR PY=2002:2005)

11/3,K/1 (Item 1 from file: 275)  
DIALOG(R)File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

02429471 SUPPLIER NUMBER: 63975291 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
A CD-ROM for Word Lovers. (Random House Webster's Unabridged  
Dictionary) (Software Review) (Evaluation)  
Pack, Thomas  
Link-Up, 17, 4, 36  
July, 2000  
DOCUMENT TYPE: Evaluation ISSN: 0739-988X LANGUAGE: English  
RECORD TYPE: Fulltext  
WORD COUNT: 1343 LINE COUNT: 00102

... up the meaning of logophile and 314,999 other words. It's Random House Webster's Unabridged Dictionary on CD-ROM, which includes content from the second edition of the printed Random House Webster's Unabridged Dictionary, first published in 1987. The electronic version, published last September, has been updated with hundreds of new...

...the biographic and geographic entries--have been updated. Random House Webster's is the only unabridged American dictionary on CD-ROM, according to the publisher, and it includes several features besides definitions:  
\* You can listen...  
...more than 120,000 audio pronunciations, including variant pronunciations of the same word.  
\* Over 2,400 entries include a link to a map or illustration.  
\* Besides searching for the main entry words, you can search for words within the definitions.  
\* Entries can be bookmarked for quick retrieval.  
\* You can print entries or copy and paste them into another document.  
\* Access to the dictionary can be integrated within WordPerfect, Microsoft Word, and other applications.  
\* You can create your own user dictionary.  
Another nice feature--especially if English isn't your native language--is that the text in the...

...and dialog boxes can be displayed in English, French, or Spanish. Overall, Random House Webster's Unabridged Dictionary on CD-ROM is an excellent resource for anyone who needs to consult a dictionary often.  
Search options  
The main interface is easy to use and somewhat intuitive. When you start typing...

11/3,K/2 (Item 2 from file: 275)  
DIALOG(R)File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

02129465 SUPPLIER NUMBER: 20080369 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
Computer, take a memo. (speech recognition and dictation software for PCs)  
(PC Tech Tutor) (Technology Information) (Column)  
Randall, Neil  
PC Magazine, v17, n1, p235(2)  
Jan 6, 1998  
DOCUMENT TYPE: Column ISSN: 0888-8507 LANGUAGE: English  
RECORD TYPE: Fulltext; Abstract  
WORD COUNT: 2294 LINE COUNT: 00186

... to do is to train it--that is, provide it with audio samples that correspond to the built-in dictionary and acoustic models. During a typical training routine, the software presents prose passages on the screen, and...

...As you finish each passage, the software stores your voice's acoustic patterns in relation to the words of the passage in context. When the training is complete the software has a reasonably detailed acoustic model of your voice; when you dictate original documents, it draws on that acoustic model by comparing your speech with the words or phrases in a context similar to that recorded during the training.

Note that if more than one person will be using...

11/3, K/3 (Item 3 from file: 275)  
DIALOG(R) File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

01356444 SUPPLIER NUMBER: 08429884 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
An AI-based approach to machine translation in Indian languages.  
(artificial intelligence)  
Raman, S.; Alwar, N.  
Communications of the ACM, v33, n5, p521(7)  
May, 1990

ISSN: 0001-0782 LANGUAGE: ENGLISH RECORD TYPE: FULLTEXT; ABSTRACT  
WORD COUNT: 5135 LINE COUNT: 00433

... As in this example, a search for a match with a query word is performed. The query dictionary, which contains all the query-related words, is used for the search, and this returns a success...

...as In. The next item in the prediction list is the agent. The search in the pronouns dictionary identifies the agent. This builds the semantic representation gradually and ends up with a phrase level representation as Figure 3 shows.

The...

...slots have a number of predefined sub-slots.  
The analyzer then identifies "Delhi" as the destination station.  
Next, when it encounters the word 'express:' it creates a prediction list and adds the train name to it. Then, using the stations dictionary, it checks whether the previous word is a valid name of a train. If it encounters a...

11/3, K/4 (Item 4 from file: 275)  
DIALOG(R) File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

01352501 SUPPLIER NUMBER: 08183498 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
Spellbound? (SpellCode programming language spell checker) (Product  
announcement)  
Kodama, David  
Data Based Advisor, v8, n1, p131(1)  
Jan, 1990  
DOCUMENT TYPE: product announcement ISSN: 0740-5200 LANGUAGE:  
ENGLISH RECORD TYPE: FULLTEXT  
WORD COUNT: 215 LINE COUNT: 00018

... text displayed to users, or it can check the entire program file. It comes with an English dictionary and a dictionary of common computer terms. SpellCode understands dBASE III PLUS, dBASE IV, FoxBASE+, FoxPro, Clipper (and most Clipper libraries), dBXL, Quicksilver, C, Pascal, BASIC, R:BASE, Paradox, and DOS file keywords.

You can build personal dictionaries for specific industries. SpellCode can check the contents of character and memo fields and understands DBF structures...

...also check Lotus and Symphony worksheets. SpellCode's interface has a

dual window that can display the context in which the words appear, suggested corrections, "fill-in-the-blanks" forms, pop-up selection windows for setting program options, and context-sensitive help screens.

SpellCode costs \$99.95 including dictionaries, support utilities, and the manual. It requires DOS 2.0 or later and 256K RAM; a hard...

11/3,K/5 (Item 5 from file: 275)  
DIALOG(R) File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

01293495 SUPPLIER NUMBER: 07172278 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
Thunder II en route to the shelves: spelling checker is modular, fast.  
(Electronic Arts Inc.) (product announcement)  
Norr, Henry  
MacWEEK, v3, n14, p14(1)  
April 4, 1989  
DOCUMENT TYPE: product announcement ISSN: 0892-8118 LANGUAGE:  
ENGLISH RECORD TYPE: FULLTEXT; ABSTRACT  
WORD COUNT: 371 LINE COUNT: 00030

... spelling checker you can buy for the Macintosh," checks approximately 100 words per second against a main dictionary of 86,000 words plus one or more user dictionaries. It installs itself automatically whenever an application or desk accessory specified in the Thunder II Control Panel...

...a user-specified beep sound any time an unrecognized word is typed.  
\* Selection checking of text already entered, with the suspect word shown in context along with the alternatives Thunder suggests.  
Thunder II detects double words and some capitalization and punctuation mistakes as well as basic spelling. It also has an abbreviation-expanding glossary and a batch search/replace function that can make multiple changes to a document in one operation...

...able to bomb it, and that's what my job is."

"It means I need only to create one set of personal dictionaries rather than several," said John Kendrick, a professor of sociology at Bucknell University in Lewisburg, Pa. "And..."

11/3,K/6 (Item 6 from file: 275)  
DIALOG(R) File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

01153331 SUPPLIER NUMBER: 00619499 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
Word Takes Another Forward Stride.  
Burns, D.; Venit, S.  
PC Magazine, v4, n13, p151-153  
June 25, 1985  
DOCUMENT TYPE: evalution ISSN: 0888-8507 LANGUAGE: ENGLISH  
RECORD TYPE: FULLTEXT; ABSTRACT  
WORD COUNT: 2228 LINE COUNT: 00169

... Spell has always been an excellent program, but this new release includes two major advances. The Standard dictionary now contains over 80,000 words, and Spell now runs directly from the Word menu.

The original Spell used a dictionary licensed from Oasis Systems, the developer of word Plus, but the new version's dictionary was created especially for Microsoft. It includes many more prefixes, suffixes, derived words, and proper names such as states...

...first saves the file you've been working on, then reads the file, and then reads the dictionary to check the spelling. When this process is completed, Spell reports back to you how many words...

...word or similar words. You can also choose to ignore the word and go on to the next one or to add the word to either the program's standard dictionary or your own dictionary.

Spell has five additional programs that can be run either directly from within...

11/3,K/7 (Item 7 from file: 275)  
DIALOG(R) File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

01153257 SUPPLIER NUMBER: 00602711 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
**All in the Family: The Perfect Components.**  
Raskin, R.; Christian, K.  
PC Magazine, v4, n10, p223-224  
May 14, 1985  
DOCUMENT TYPE: evaluton ISSN: 0888-8507 LANGUAGE: ENGLISH  
RECORD TYPE: FULLTEXT; ABSTRACT  
WORD COUNT: 4865 LINE COUNT: 00376

... your way through the list, you can instruct PS to ignore a word, add it to the **dictionary**, or mark it in the text. Then, when Perfect Writer is started, it works through the misspelled words in context.

When you add words to the dictionary supplied with Perfect Speller, you increase the chance that the program will miss an incorrect word. Therefore, rather than adding a few hundred words to the Speller dictionary, it is better to create your own dictionaries for specialized topics. However, PS searches only the dictionary each time it is invoked. It would be preferable if it could use the main dictionary in conjunction with a customized dictionary. Dual dictionaries make it easier to tell a spelling checker about technical terms without adulterating the main dictionary.

Perfect Writer's literature boasts that the Speller uses a 50,000-word **dictionary**. Perhaps they have invented sub-bit storage methods. Their **dictionary** file is 22,000 bytes long. That works out to about 4 bits for each word. Evenwith the best compression techniques, a 50,000-word **dictionary** should require over 100,000 bytes of storage.

Perfect Thesaurus

An interesting extra in the Writer package...

11/3,K/8 (Item 8 from file: 275)  
DIALOG(R) File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

01104081 SUPPLIER NUMBER: 00590265 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
**'Correct Star.'**  
Kennedy, D.  
PC Week, v1, n30, p55-56  
July 31, 1984  
DOCUMENT TYPE: evaluton ISSN: 0740-1604 LANGUAGE: ENGLISH  
RECORD TYPE: FULLTEXT; ABSTRACT  
WORD COUNT: 1153 LINE COUNT: 00089

... automatic right justification, soft hyphenation and some "housekeeping" matters, such as designating the drives where the various dictionaries and files reside.

When the spelling checker is started, CorrectStar loads an internal dictionary into RAM. It...

...the main dictionary, but which a user expects will appear frequently in his or her documents. The **personal dictionary** can contain as many as 1,500 words, or about 18K bytes.

Many **personal dictionaries** can be created and stored on the same disk, provided there is sufficient storage space. Only one can be used

...keystroke, choose one of six options: Correct as suggested, correct globally (throughout the document), look at the **next** of previous suggestion, add the word to the **personal dictionary** so it will remain a permanent entry, bypass the word in this one instance or ignore the...

...types "hear and there" will not be corrected since the misspelled words are, in fact, correctly spelled **words** in a different **context**.

However, CorrectStar's ability to catch suspect words and suggest alternatives is impressive. For example, anyone who...

11/3,K/9 (Item 9 from file: 275)  
DIALOG(R) File 275:Gale Group Computer DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

01022444 SUPPLIER NUMBER: 00520613  
The Random House Proofreader.  
Meilach, D.Z.  
Interface Age, v8, n5, p24  
May, 1983  
ISSN: 0147-2992 LANGUAGE: ENGLISH RECORD TYPE: ABSTRACT

...ABSTRACT: and can be used with Word Star and other word processing programs. This program will display misspelled **words** in **context**, list misspelled **words**, store a custom document **dictionary** and provide **dictionary** reference help for spelling correction. Program weaknesses include the **dictionary** help choices and some error detection features. Samples of the initial menu, word- flagging and the options...

11/3,K/10 (Item 1 from file: 621)  
DIALOG(R) File 621:Gale Group New Prod.Annou. (R)  
(c) 2005 The Gale Group. All rts. reserv.

01175091 Supplier Number: 42413777 (USE FORMAT 7 FOR FULLTEXT)  
MS-DOS FILE CONVERSION HIGHLIGHTS FEATURES OF PANASONIC'S 1991 DESKTOP WORD PROCESSORS  
News Release, p1  
Oct 2, 1991  
Language: English Record Type: Fulltext  
Document Type: Magazine/Journal; Trade  
Word Count: 665

... ever to operate, Word Count, and Double Word Count, which identifies double words in the text. Features include insert, delete, word search, next word search, word replace, block move/delete/ copy , word wrap, format memory, automatic pagination and much more.

In addition, both offer a number of other...

...made with a touch of the button. Up to 120 words can also be added to a built -in user **dictionary** and up to 2,000 characters can be stored in Phrase Memory.

(2) Spelling programs developed and copyrighted by Houghton Mifflin Company, publishers of The American Heritage **Dictionary**.

Both word processors also feature built -in Thesaurus (3), with approximately 45,000 reference words and 500,000 synonyms. Move the cursor to...

11/3,K/11 (Item 2 from file: 621)  
DIALOG(R)File 621:Gale Group New Prod.Annou.(R)  
(c) 2005 The Gale Group. All rts. reserv.

01145742 Supplier Number: 41650257 (USE FORMAT 7 FOR FULLTEXT)  
**JAKE GELLER'S SPELL CHECKING ENGINE (TM) ADDS SPELL CHECKING TO CLIPPER AND C APPLICATIONS**  
News Release, pl  
Nov 1, 1990  
Language: English Record Type: Fulltext  
Document Type: Magazine/Journal; Trade  
Word Count: 937

... be a  
seamless part of existing applications."

The Spell Checking Engine is supplied with an English language dictionary containing over 100,000 words. The dictionary is stored in a compressed format to minimize disk storage requirements and enhance lookup speed.

A maintenance utility is provided which allows adding or removing custom words or abbreviations to the main dictionary. This is important to applications in medical, legal, scientific and other specialized fields.

Auxiliary dictionaries can also be created for this purpose. Users can add their own words "on the fly," and these additions can be incorporated into the main or auxiliary dictionaries if desired.

The Spell Checking Engine provides flexibility to accommodate a broad spectrum of end-user computing...

...optimized for speed, memory utilization or both, permitting spell checking even in low memory situations.

The dictionary can be loaded into memory for optimal performance if sufficient free memory is available, or it can...

...manner the developer chooses, e.g., popup box, bar list, etc. Suggestions can be tailored to the context of the word, including such factors as capitalization. Unique features allow checking over an expanded alphabet, including alphabetic, numeric and special characters.

Jake Geller's Spell Checking Engine sells for \$249, including all dictionaries, support utilities and an extensive user manual. This price includes 25 licenses to distribute complete applications containing...

11/3,K/12 (Item 1 from file: 636)  
DIALOG(R)File 636:Gale Group Newsletter DB(TM)  
(c) 2005 The Gale Group. All rts. reserv.

02577212 Supplier Number: 45205057 (USE FORMAT 7 FOR FULLTEXT)  
**VOICE PROCESSING CORP**  
Voice Technology News, v6, n25, pN/A  
Dec 13, 1994

Language: English Record Type: Fulltext  
Document Type: Magazine/Journal; Trade  
Word Count: 126

(USE FORMAT 7 FOR FULLTEXT)  
TEXT:  
...because it can obtain the pronunciation of words used in applications from any combination of VProFlex's built -in 100,000 word phonetic dictionary of acoustic word models, text-to-speech phonetic output using industry standard phonetic alphabets and application supplied "pronounce" statements. Other enhancements include word spotting, application specific context -free grammar and continuous input. VProFlex evaluation kits are available for qualified developers and a full software...

11/3,K/13 (Item 1 from file: 16)  
DIALOG(R)File 16:Gale Group PROMT(R)  
(c) 2005 The Gale Group. All rts. reserv.

08913540 Supplier Number: 77275800 (USE FORMAT 7 FOR FULLTEXT)  
Visa's Curriculum Site Brings Praise, Questions.(educational website supposedly promotes use of Visa credit cards)(Brief Article)  
Kingston, Jennifer A.; Kuykendall, Lavonne  
American Banker, v166, n159, p9  
August 17, 2001  
Language: English Record Type: Fulltext  
Article Type: Brief Article  
Document Type: Magazine/Journal; Trade  
Word Count: 856

... to educators.  
According to the Web site, teachers praised [www.practicalmoneyskills.com](http://www.practicalmoneyskills.com) for, among other things, its glossary of banking and financial terms, its detailed lesson plans for preschool through 12th grade, and for "not..."

...curriculum does emphasize using the term "check card" -- which is Visa's particular name for a generic product , the debit card. In the glossary , there is a detailed definition of the term "check card" (which does not include the word debit). Next to the term "debit card," it says, "See 'check card.' "

The curriculum includes a 10-page lesson on shopping for...

11/3,K/14 (Item 2 from file: 16)  
DIALOG(R)File 16:Gale Group PROMT(R)  
(c) 2005 The Gale Group. All rts. reserv.

06307257 Supplier Number: 54521199 (USE FORMAT 7 FOR FULLTEXT)  
Word Processing.(Microsoft Word)(Technology Tutorial)  
Campbell, George  
PC World, v17, n4, p289(1)  
April, 1999  
Language: English Record Type: Fulltext  
Document Type: Magazine/Journal; General Trade  
Word Count: 1416

... in the Macros list.  
4. Select &File from the Change What Menu drop-down list.  
5. Click Add , then click Close.  
Fix Errors In Word's Custom Dictionaries  
It happens to almost everyone: While rushing through a spelling check, you mistakenly add a misspelled (or mistyped) word to Word's custom dictionary . From then on, Word doesn't flag that misspelled word as incorrect, so it may go unnoticed. Fortunately, the custom dictionary

is a simple text document that you can easily edit. Here's how:

1. Select Tools, Options...

...6 and 7, click the Spelling tab.

3. In Word 97, click Dictionaries. In Word 7, click Custom Dictionaries. Word 6 users can skip this step.

4. Select CUSTOM.DIC or the dictionary you want to change in the Custom Dictionaries list.

5. Click Edit, then click OK in the advisory dialog box that appears.

6. After Word...

...you added erroneously.

7. Once your edits are finished, select File, Save, then File, Close.

8. In Word 7 and 97, you'll need to select Tools, Options, click the Spelling & Grammar tab, and enable...

...prior to your edits. This is necessary because Word turns this feature off when you edit a custom dictionary.

9. You can also add words to your custom dictionary if you want. Just type them in during Step 6 and save.

Drag and Copy Text In...

11/3,K/15 (Item 3 from file: 16)

DIALOG(R)File 16:Gale Group PROMT(R)

(c) 2005 The Gale Group. All rts. reserv.

03153221 Supplier Number: 44306420  
Software firm automates Japanese translations

Mass High Tech, p18

Dec 20, 1993

Language: English Record Type: Abstract

Document Type: Magazine/Journal; Trade

ABSTRACT:

...tems of the probability of being correct. The program also learns the preferred usage of sentences or words in a certain context. Users may add all or parts of 19 separate special dictionaries for physics, chemistry, engineering, business, architecture or other applied sciences. The program also permits users to create custom dictionaries.

11/3,K/16 (Item 4 from file: 16)

DIALOG(R)File 16:Gale Group PROMT(R)

(c) 2005 The Gale Group. All rts. reserv.

01213339 Supplier Number: 41397407

PANASONIC'S NEW LAPTOP WORD PROCESSOR IS SMALL IN SIZE ONLY

News Release, pl

June 21, 1990

Language: English Record Type: Abstract

Document Type: Magazine/Journal; Trade

ABSTRACT:

...word processor that brings new meaning to the word "personal." About the size of a standard hardcover dictionary, the KS-WL50 offers exceptional portability, and is perfect for use by travelling executives, by college students...

...the KX-WL50 has full word processing capabilities built-in, including word search and replace, block move/ copy /delete, next word , word wrap, format memory and memory quantity. The unit also features Panasonic's 63,000-word Accu-Spell Plus (TM) built -in dictionary , with 120 words user-programmable. ...

11/3,K/17 (Item 5 from file: 16)  
DIALOG(R)File 16:Gale Group PROMT(R)  
(c) 2005 The Gale Group. All rts. reserv.

01006827 Supplier Number: 41098820 (USE FORMAT 7 FOR FULLTEXT)  
Canon's introduction at Winter Consumer Electronics Show next week will  
include 2 word processors, 2 personal copiers and fax phone  
Consumer Electronics, pN/A  
Jan 1, 1990  
Language: English Record Type: Fulltext  
Document Type: Newsletter; Trade  
Word Count: 103

(USE FORMAT 7 FOR FULLTEXT)  
TEXT:  
Canon's introductions at Winter Consumer Electronics Show next week will  
include 2 word processors, 2 personal copiers and fax phone, all aimed  
at home office market. Canon BW 70 word...

...LCD, 3.5", 720 K-byte floppy disc, capability for more than 20  
languages, 90,000-word built -in dictionary . Other model, TW 40 (\$700 in  
Oct.), has same features but slower printing speed. Personal copiers, at...

11/3,K/18 (Item 1 from file: 148)  
DIALOG(R)File 148:Gale Group Trade & Industry DB  
(c)2005 The Gale Group. All rts. reserv.

09049416 SUPPLIER NUMBER: 18789658 (USE FORMAT 7 OR 9 FOR FULL TEXT)  
Editing Word's dictionaries. (tips for using Microsoft's Word word  
processing software) (question-and-answer) (Product Support) (Brief  
Article)  
PC World, v14, n11, p302(1)  
Nov, 1996  
DOCUMENT TYPE: Brief Article ISSN: 0737-8939 LANGUAGE: English  
RECORD TYPE: Fulltext  
WORD COUNT: 151 LINE COUNT: 00014

Li Yu, People's Republic of China  
Word stores its custom dictionary as an easy-to-edit ASCII file.  
Select Tools\*Options, then click the Spelling tab in the Options dialog  
box. Choose the dictionary you want to edit from the Custom  
Dictionaries list and click Edit. Click Yes in the resulting dialog box,  
and click OK if asked to...

...in the Options dialog box to close it, then make your changes in  
custom.dic. Select File\* Save , then File\* Close .  
In Word 7, click the Custom Dictionaries button in the Options  
dialog box before choosing the dictionary you want to edit.

11/3,K/19 (Item 1 from file: 15)  
DIALOG(R)File 15:ABI/Inform(R)  
(c) 2005 ProQuest Info&Learning. All rts. reserv.

02912078 860468781  
Making Its Own Bed  
Stankevich, Debby Garbato  
Retail Merchandiser v40n5 PP: 83-85 May 2000  
ISSN: 1530-8154 JRNL CODE: DMD  
WORD COUNT: 2770

...TEXT: aggressive- including Wal-Mart, ShopKo and Kmart-Bed Bath & Beyond  
is not trying to put its whole store online tomorrow. "Our next step is

to get **word** out to the 240 **stores** , " says Eisenberg. "We haven't seen it show up on shopping bags yet. We haven't put..."

...m. will have the finer points of anodized aluminum (cookware) explained. The customer rep even has a **product guide glossary** to assist."

Unlike in its brick-and-mortar operation, the company is using a fulfillment center to...

11/3,K/20 (Item 2 from file: 15)  
DIALOG(R) File 15:ABI/Inform(R)  
(c) 2005 ProQuest Info&Learning. All rts. reserv.

02791469 696196021  
**creative techniques**  
Bajaj, Geetesh  
Presentations v18n9 PP: 16-17 Sep 2004  
ISSN: 1072-7531 JRNL CODE: PRS  
WORD COUNT: 1302

...TEXT: taskpane metaphor to spawn new capabilities. By default, the Research task pane can be used as a **dictionary** and thesaurus, and includes access to reference books and research sites. You can add even more reference...

...2 The Research task pane produces a list of definitions from any of the several built-in **dictionary** versions.

A defining resource

Most PowerPoint users I surveyed turn to the Research task pane as a...

...version of the North America Encarta Dictionary. For example, the Research task pane gave me six detailed **definitions** of the **word** visible, along with a text pronunciation (see EXAMPLE 2). You can also switch to the United Kingdom version of the Encarta **Dictionary**, which showed me seven **definitions** for the same **word**. A search for the **word** technology got me three **definitions** with sample sentences using the **word** in all **contexts**. Several noun, adjective and adverb forms of the **word** were also displayed. Unfortunately, the task pane has no audio component to let you **hear** how a **word** is pronounced.

The Research task pane also **includes** several built-in thesauri. Searching within this pane, I found three English, two French and two Spanish...

11/3,K/21 (Item 1 from file: 647)  
DIALOG(R) File 647:cmp Computer Fulltext  
(c) 2005 CMP Media, LLC. All rts. reserv.

01145690 CMP ACCESSION NUMBER: WIN19971115S0013  
**Applications - From contact managers to spreadsheets and databases, these tips will help you stay organized and boost your productivity.**  
WINDOWS MAGAZINE, 1997, n 811A, PG101  
PUBLICATION DATE: 971115  
JOURNAL CODE: WIN LANGUAGE: English  
RECORD TYPE: Fulltext  
SECTION HEADING: 2,001 Tips  
WORD COUNT: 21282

... the paragraph style or styles you don't need. Click on Delete, Yes and Close.

Add/Use **Glossary Entries**

Store frequently used text in a **glossary** entry. To create an

entry, highlight the desired text and select Edit/Glossary. Type a name in the Glossary Entry Name box and click on **Create**. To use a **glossary** entry, place the cursor where you want the text and select Edit/ **Glossary**. Choose the appropriate **glossary** entry name from the **Glossary** Entry Name box and click on **Insert**.

Word Pro 97

TASK....SHORTCUT

Delete **next word** ....Ctrl+Del

Delete **previous word** ....Ctrl+Backspace

Delete a row in a table....Ctrl+- (minus key on numeric keypad)

Center text....Ctrl...

...Alt+

Go To....Ctrl+

Insert a row in a table....Ctrl++(plus key on numeric keypad)

Insert **glossary** record....Ctrl+

Redo action....Alt+Shift+Backspace or Ctrl+Shift+

Show/Hide set of SmartIcons....Ctrl+

Undo...

11/3, K/22 (Item 1 from file: 610)

DIALOG(R)File 610:Business Wire

(c) 2005 Business Wire. All rts. reserv.

00524722 20010522142B4833 (USE FORMAT 7 FOR FULLTEXT)

WordLogic Corporation Launches WordLogic Text Input Software for Wireless Information Devices at VORTEX 2001

Business Wire

Tuesday, May 22, 2001 09:08 EDT

JOURNAL CODE: BW LANGUAGE: ENGLISH RECORD TYPE: FULLTEXT

DOCUMENT TYPE: NEWSWIRE

WORD COUNT: 613

...is displayed onscreen and is compact and intuitive. With a tap of a letter, it predicts the **next** most probable letters and **words**. In two taps, most words can be **entered**. More complex words are completed with WordLogic Corporation's proprietary WordChunking(TM) technology. The WordLogic keyboard remembers common phrases and adapts to each user's vocabulary incorporating both standard and **custom dictionaries**.

File 348:EUROPEAN PATENTS 1978-2005/Jun W04

(c) 2005 European Patent Office

File 349:PCT FULLTEXT 1979-2005/UB=20050707,UT=20050630

(c) 2005 WIPO/Univentio

Set	Items	Description
S1	17214	DICTIONAR??? OR GLOSSAR???
S2	783719	WORD? ? OR KEYWORD? ? OR TERM? ?
S3	91970	S2 (5N) (DEFIN??? OR DEFINITION? ? OR MEANING)
S4	25234	S2 (5N) CONTEXT???
S5	14389	(ADJACENT OR NEXT OR NEAR OR NEARBY OR CLOSE OR AROUND OR - SAME () SENTENCE OR PROXIMAL? OR PROXIMITY OR VICINIT??? OR ADJ- OIN??? OR NEIGHBOR??? OR BESIDE? ? OR SURROUND???) (5W) S2
S6	28141	(CONTEXT OR S5) (7N) (STOR??? OR SAV??? OR ADD??? OR INSERT?- ?? OR INCORPORAT??? OR INCLUD??? OR SUBMIT???? OR ENTER??? OR COPY??? OR COPIE? ?)
S7	925	(S1 OR S3) (50N) S4:S5 (50N) S6
S8	327	(S1 OR S3) (50N) S4:S5 (50N) S6 (50N) DEFINITION? ?
S9	47	S8 AND IC=G06F
S10	1560	S1 (5N) (BUILD??? OR BUILT OR CONSTRUCT??? OR CREAT??? OR GE- NERAT??? OR PRODUC??? OR ESTABLISH?)
S11	368	(PERSONAL OR CUSTOM? OR ALTERNATE OR ALTERNATIVE OR USER OR AUXILIARY) (2W) S1
S12	26	(S1 OR S3) (50N) S4:S5 (50N) S6 (50N) S10:S11
S13	4899	CONTEXT(5N) (STOR??? OR SAV??? OR ADD??? OR INSERT???)
S14	25	S1 (5N) S13

9/3,K/6 (Item 6 from file: 348)  
DIALOG(R)File 348:EUROPEAN PATENTS  
(c) 2005 European Patent Office. All rts. reserv.

01641369

DOCUMENT KNOWLEDGE MANAGEMENT APPARATUS AND METHOD  
DOKUMENTWISSENSVERWALTUNGS VORRICHTUNG UND VERFAHREN  
APPAREIL ET PROCEDE PERMETTANT DE TRAITER DES CONNAISSANCES DANS DES  
DOCUMENTS

PATENT ASSIGNEE:

Celestar Lexico-Sciences, Inc., (4432260), Makuhari Techno Garden D17,  
1-3, Nakase, Mihamaku, Chiba-shi, Chiba 261-8501, (JP), (Applicant  
designated States: all)

INVENTOR:

NITTA, Kiyoshi, c/o Celestar Lexico-Sciences Inc., Makuhari Techno Garden  
D17, 1-3, Nakase, Mihamaku, Chiba-shi, Chiba 261-8501, (JP)  
DOI, Hiroyumi, c/o Celestar Lexico-Sciences, Inc., Makuhari Techno Garden  
D17, 1-3, Nakase, Mihamaku, Chiba-shi, Chiba 261-8501, (JP)  
KIKUCHI, Yasuhiro, c/o Celestar Lexico-Scs, Inc., Makuhari Techno Garden  
D17, 1-3, Nakase, Mihamaku, Chiba-shi, Chiba 261-8501, (JP)  
HORAI, Hisayuki, c/o Celestar Lexico-Sciences Inc., Makuhari Techno  
Garden D17, 1-3, Nakase, Mihamaku, Chiba-shi, Chiba 261-8501, (JP)

LEGAL REPRESENTATIVE:

HOFFMANN - EITLE (101511), Patent- und Rechtsanwalte Arabellastrasse 4,  
81925 Munchen, (DE)

PATENT (CC, No, Kind, Date): EP 1473639 A1 041103 (Basic)  
WO 2003067471 030814

APPLICATION (CC, No, Date): EP 2003705045 030204; WO 2003JP1120 030204  
PRIORITY (CC, No, Date): JP 200227446 020204; JP 2002237687 020816; JP  
2002237688 020816; JP 2002237689 020816; JP 2002375457 021225

DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;  
HU; IE; IT; LI; LU; MC; NL; PT; SE; SI; SK; TR

EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO

INTERNATIONAL PATENT CLASS: G06F-017/30 ; G06F-019/00

ABSTRACT WORD COUNT: 132

NOTE:

Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; Japanese  
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200445	10891
SPEC A	(English)	200445	82558
Total word count - document A			93449
Total word count - document B			0
Total word count - documents A + B			93449

INTERNATIONAL PATENT CLASS: G06F-017/30 ...

... G06F-019/00

...SPECIFICATION semantic dictionary creating step. According to the  
present program, each of the terms (keywords, abbreviations, synonyms,  
related words , etc. included in the dictionary) entered in the  
pre-existing dictionary information is assessed, based on the term, as  
being a canonical form...

...A dictionary information processing program according to still another  
aspect of the present invention: wherein the semantic dictionary  
creating step further comprises a Web term appraising step that assesses,  
based on terms entered in a pre-existing dictionary information,  
whether each of the terms in the Web information is to be considered as a  
canonical form, variant form, or a term that is not to be used, and  
creates the semantic dictionary information from each term of the Web  
information, based on an appraisal result of the Web term appraising  
step.

This is a more specific explanation of the semantic dictionary creating step. According to the present program, each of the terms in pre-existing Web information (including...

9/3,K/7 (Item 7 from file: 348)  
DIALOG(R) File 348:EUROPEAN PATENTS  
(c) 2005 European Patent Office. All rts. reserv.

01507255

Database model, tools and methods for organizing information across external information objects  
Datenbankmodell, Werkzeuge und Verfahren zum Organisieren von Information unter externen Informationsobjekten  
Modele de base de donnees, outils et methode pour organiser des informations parmi des objets informatifs externes

PATENT ASSIGNEE:

Agilent Technologies Inc, (2885681), 395 Page Mill Road, Palo Alto, CA 94306-2024, (US), (Applicant designated States: all)

INVENTOR:

KUCHINSKY, Allan, 305 Yerba Buena Avenue, San Francisco, CA 94127, (US)  
MOH, David, 282 Romain Street, San Francisco, CA 94131, (US)  
GRAHAM, Katherine D., 6167 Silberman Drive, San Jose, CA 95120, (US)  
CREECH, Michael L., 1734 Austin Avenue, Los Altos, CA 94024, (US)

LEGAL REPRESENTATIVE:

Schoppe, Fritz, Dipl.-Ing. (55464), Patentanwalte Schoppe, Zimmermann, Stockeler & Zinkler, Postfach 71 08 67, 81458 Munchen, (DE)

PATENT (CC, No, Kind, Date): EP 1260918 A2 021127 (Basic)

APPLICATION (CC, No, Date): EP 2002011256 020522;

PRIORITY (CC, No, Date): US 863115 010522

DESIGNATED STATES: AT; BE; CH; CY; DE; DK; ES; FI; FR; GB; GR; IE; IT; LI; LU; MC; NL; PT; SE; TR

EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI

INTERNATIONAL PATENT CLASS: G06F-017/30

ABSTRACT WORD COUNT: 187

NOTE:

Figure number on first page: 1

LANGUAGE (Publication,Procedural,Application): English; English; English  
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200248	976
SPEC A	(English)	200248	12276
Total word count - document A			13252
Total word count - document B			0
Total word count - documents A + B			13252

INTERNATIONAL PATENT CLASS: G06F-017/30

...SPECIFICATION limits, ranges excluding either or both of those included limits are also included in the invention.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although...

...noted that as used herein and in the appended claims, the singular forms "a", "and", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a viewer" includes a plurality of such viewers and...

...publication provided may be different from the actual publication dates which may need to be independently confirmed.

DEFINITIONS

The term "activation" refers to enhancement of the effects of a

biological agent or stimulation of a biological or chemical process, for example.

The term "alternative" when used in the context of describing a biological story, refers to one choice among a number of possible explanations (or hypotheses) for a biological phenomenon.

The...

...note that may be associated with any item, collection, story element, diagram node, or diagram interaction.

The term "biological story" defines a high-level description or explanation of a complex biological process, formulated by a researcher, for example...

9/3,K/9 (Item 9 from file: 348)

DIALOG(R) File 348:EUROPEAN PATENTS

(c) 2005 European Patent Office. All rts. reserv.

00885417

Document display system and electronic dictionary

Dokumentanzeigesystem und elektronisches Wörterbuch

Système d'affichage de documents et dictionnaire électronique

PATENT ASSIGNEE:

Oki Electric Industry Co., Ltd., (225692), 7-12, Toranomon 1-chome  
Minato-ku, Tokyo, (JP), (Proprietor designated states: all)

INVENTOR:

Ikeno, Atsushi, Oki Electronic Industry Co.,Ltd., 7-12,Toranomon,  
1-chome, Minato-ku, Tokyo, (JP)

LEGAL REPRESENTATIVE:

Kirschner, Klaus Dieter, Dipl.-Phys. (6506), Schneiders & Behrendt  
Rechtsanwalte - Patentanwalte Sollner Strasse 38, 81479 München, (DE)

PATENT (CC, No, Kind, Date): EP 810534 A2 971203 (Basic)

EP 810534 A3 000322

EP 810534 B1 030319

APPLICATION (CC, No, Date): EP 97107176 970429;

PRIORITY (CC, No, Date): JP 96117415 960513; JP 96117661 960513; JP  
96118766 960514; JP 96118795 960514

DESIGNATED STATES: DE; FR; GB

INTERNATIONAL PATENT CLASS: G06F-017/30

ABSTRACT WORD COUNT: 78

NOTE:

Figure number on first page: NONE

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available	Text	Language	Update	Word Count
	CLAIMS A	(English)	199711W4	2370
	CLAIMS B	(English)	200312	2418
	CLAIMS B	(German)	200312	2260
	CLAIMS B	(French)	200312	2773
	SPEC A	(English)	199711W4	18752
	SPEC B	(English)	200312	18961
Total word count	- document A			21126
Total word count	- document B			26412
Total word count	- documents A + B			47538

INTERNATIONAL PATENT CLASS: G06F-017/30

...SPECIFICATION link to the "Global Slogan" document, he will see first the display in FIG. 13 (without a dictionary mode button, because the "Corporate Guidance" document is a hypertext document), then the display in FIG. 14 (with a dictionary mode button, because the "Global Slogan" document is not a hypertext document).

If the user selects the dictionary mode button on the display in FIG. 14, the linked document server 2 activates the dictionary linker 4,

which generates a result file according to the look-up counts maintained by the **dictionary** access tabulator 18. The linked document server 2 adds an ordinary mode button, and sends this file...

...ten are used as described above. The first line in FIG. 61 is the ordinary mode button added by the linked document server 2. The **next** line contains only the word "We," because the look-up count (one) for this word exceeds the threshold value (zero) for pronouns. The **next** line contains the word "draw" together with a **dictionary** look-up command tag and closing tag, because the look-up count (three) for this word does not exceed the threshold value (ten) for verbs. Other lines are generated similarly by the **dictionary** linker 4. The client device 1 displays this result file as shown in FIG. 62.

The user...

...will obtain a Japanese definition as shown in FIG. 16 or 27. If the user selects the word "We," however, no definition will be returned and the display in FIG. 62 will remain unchanged, because no **dictionary** access tag is attached to this word.

By not underlining words that the user has already looked...

...SPECIFICATION link to the "Global Slogan" document, he will see first the display in FIG. 13 (without a **dictionary** mode button, because the "Corporate Guidance" document is a hypertext document), then the display in FIG. 14 (with a **dictionary** mode button, because the "Global Slogan" document is not a hypertext document).

If the user selects the **dictionary** mode button on the display in FIG. 14, the linked document server 2 activates the **dictionary** linker 4, which generates a result file according to the look-up counts maintained by the **dictionary** access tabulator 18. The linked document server 2 adds an ordinary mode button, and sends this file...

...ten are used as described above. The first line in FIG. 61 is the ordinary mode button added by the linked document server 2. The **next** line contains only the word "We," because the look-up count (one) for this word exceeds the threshold value (zero) for pronouns. The **next** line contains the word "draw" together with a **dictionary** look-up command tag and closing tag, because the look-up count (three) for this word does not exceed the threshold value (ten) for verbs. Other lines are generated similarly by the **dictionary** linker 4. The client device 1 displays this result file as shown in FIG. 62.

The user...

...will obtain a Japanese definition as shown in FIG. 16 or 27. If the user selects the word "We," however, no definition will be returned and the display in FIG. 62 will remain unchanged, because no **dictionary** access tag is attached to this word.

By not underlining words that the user has already Looked...

9/3,K/17 (Item 2 from file: 349)  
DIALOG(R) File 349:PCT FULLTEXT  
(c) 2005 WIPO/Univentio. All rts. reserv.

01214770 \*\*Image available\*\*

METHOD AND APPARATUS FOR LANGUAGE PROCESSING  
TECHNIQUE ET DISPOSITIF DE TRAITEMENT LINGUISTIQUE

Patent Applicant/Assignee:

WHITE SMOKE INC, P.O. Box 391, The Brandywine Building, 1000 West Street,  
17th Floor, Wilmington, Delaware 19801, US, US (Residence), US  
(Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

OVIL Joel, 26th Avenue, Houghton, Johannesburg, ZA, ZA (Residence), IL  
(Nationality), (Designated only for: US)

BRENER Liran, 12 Hagula Street, 45272 Hod Hasharon, IL, IL (Residence),  
IL (Nationality), (Designated only for: US)

Legal Representative:

SOCKOL Marc A (et al) (agent), Squire, Sanders & Dempsey L.L.P., 600  
Hansen Way, Palo Alto, California 94304-1043, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200522294 A2 20050310 (WO 0522294)

Application: WO 2004US21779 20040706 (PCT/WO US04021779)

Priority Application: US 2003613146 20030703

Designated States:

(All protection types applied unless otherwise stated - for applications  
2004+)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM  
DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC  
LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NI NO NZ OM PG PH PL PT RO  
RU SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW  
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PL PT RO  
SE SI SK TR

(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG

(AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 12129

Main International Patent Class: G06F

Fulltext Availability:

Detailed Description

Detailed Description

... of author, and for each of a plurality of sentences within the training text: identifying pairs of words, W1 and W2, with known contexts within a sentence, used together in conjunction, and designating matches between...

...yet further provided in accordance with a preferred embodiment of the present invention a method for resolving context ambiguity within a natural language sentence, including providing a plurality of context @5 -equivalon-co gr9LiR&

Hith

@p ip e 1 6 s  
q!@jf  
pai s of the...

...that are used in the same context, parsing a natural language sentence to identify grammatical types of words within the sentence, identifying context equivalence groups to which words within the sentence !O belong, a-ad resolving contexts of ambiguous words within the sentence, consistent with matches between the identified context equivalence groups,

There is additionally provided in accordance with a preferred embodiment of the present- invention apparatus for resolving context ambiguity within a natural language sentence, including a memory for storing a plurality of i5 context equivalence groups, with specific pairs of the context equivalence groups designated as being matched, a context equivalence...

...language sentence to identify grammatical types of WO 2005/022294 PCT/US2004/021779 sentence belong, and resolving contexts of ambiguous words within the sentence consistent with matches between the identified context equivalence groups.

The following definitions are employed throughout the specification and claims.

- 1 . Ambigg - more than one possible meaning for a word
2. Context Equivalence Group, also Group - a group of words of a common Grammatical Type that can be used...

9/3, K/20 (Item 5 from file: 349)

DIALOG(R) File 349:PCT FULLTEXT  
(c) 2005 WIPO/Univentio. All rights reserved.

01153716 \*\*Image available\*\*

SEMANTIC KNOWLEDGE RETRIEVAL MANAGEMENT AND PRESENTATION  
SYSTEME ET PROCEDE POUR UNE EXTRACTION, UNE GESTION, UNE CAPTURE, UN  
PARTAGE, UNE DECOUVERTE, UNE DISTRIBUTION ET UNE PRESENTATION DE  
CONNAISSANCES SEMANTIQUES

Patent Applicant/Assignee:

NERVANA INC, 10838 Main Street, Bellevue WA, 98004, US, US (Residence),  
US (Nationality)

Inventor(s):

OMOIGUI Nosa, 549 239th Avenue S.E., Redmond, WA 98074, US,

Legal Representative:

BLACK Richard T (agent), Black Lowe & Graham PLLC, 816 Second Avenue,  
Seattle, WA 98104, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200475466 A2-A3 20040902 (WO 0475466)

Application: WO 2004US4674 20040217 (PCT/WO US04004674)

Priority Application: US 2003447736 20030214

Designated States:

(All protection types applied unless otherwise stated - for applications  
2004+)

AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM  
DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC  
LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NI NO NZ OM PG PH PL PT RO  
RU SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW  
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE  
SI SK TR  
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG  
(AP) BW GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW  
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 160617

Main International Patent Class: G06F-017/30

Fulltext Availability:

Detailed Description

Detailed Description

... Conversations

....., 146 3. Semantic Thread

Management

....., 147 Q. SAMPLE SCREEN SHOTS

148

R. SPECIFICATION FOR SEMANTIC QUERY DEFINITIONS & VISUALIZATIONS FOR  
THE  
INFORMATION NERVOUS SYSTEM

....., 148 1 Semantic Images & Motion

....., 148 2. The Smart

Hourglass

....., 153...

...with particular words, or particular strings of several sentences or paragraphs.

The reason for this is that **words** do not denote or connote **meaning** one to one as, for example, numerals tend to do. Put differently, certain meanings can be denoted or connoted by several different **words** or an essentially infinite combination of **words**, and, conversely, certain

9

words or combinations of words can denote or connote several different meanings. Despite this infinite many-to-many network of possibilities human beings can isolate (because of **context**, experience, reasoning, inference, deduction, judgment, learning and the like) isolate probable meanings, at least tolerably effectively most...

9/3, K/23 (Item 8 from file: 349)  
DIALOG(R) File 349:PCT FULLTEXT  
(c) 2005 WIPO/Univentio. All rts. reserv.

00883010 \*\*Image available\*\*  
DYNAMIC PERSONALIZATION METHOD OF CREATING PERSONALIZED USER PROFILES FOR  
SEARCHING A DATABASE OF INFORMATION  
PROCEDE DE PERSONNALISATION DYNAMIQUE DE CREATION DE PROFILS UTILISATEUR  
PERSONNALISES, AUX FINS DE RECHERCHE DANS UNE BASE DE DONNEES  
D'INFORMATIONS

Patent Applicant/Assignee:

SEASEER R & D LLC, P.O. Box 1213, Kula, Maui, HI 96790, US, US  
(Residence), US (Nationality)

Inventor(s):

FABLES Wylci, 214 Kawehi Place, Kula, Maui, HI 96790, US,  
PARK Jore, 214 Kawehi Place, Kula, Maui, HI 96790, US,  
COLT Jonathan, P.O. Box 1630, Makawao, Maui, HI 96768, US,

Legal Representative:

CHONG Leighton K (agent), Ostrager Chong & Flaherty (Hawaii), Suite 1200,  
841 Bishop Street, Honolulu, HI 96813, US

Patent and Priority Information (Country, Number, Date):

Patent: WO 200217145 A1 20020228 (WO 0217145)  
Application: WO 2001US26574 20010824 (PCT/WO US0126574)  
Priority Application: US 2000228154 20000825; US 2001278 20010823

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ  
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR  
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL  
TJ TM TR TT TZ UA UG UZ VN YU ZA ZW  
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR  
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG  
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW  
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 8223

Main International Patent Class: G06F-017/30

Fulltext Availability:

Detailed Description

Detailed Description

... an extremely fine-grained metric. Typically a database in today's marketing world has 8 - 16 demographic **definitions**. Now that more powerful machines are coming on line up to several hundred demographic categories are sometimes seen. With a typical database structure, adding a demographic **definition** directly affects performance measured in

response time. The user profiling method allows for an actual and "able to be applied" demographic group of one definition per user or multiple millions without affecting performance. The similarities and functional aspects of the individual demographic definitions can easily be applied or associated with others in the group creating self-defining micro-demographic pathways or groups. Traditional demographic understanding can be extracted and applied within this process by defining groups of users using words, thereby creating a built in transition to legacy systems of consumer data, and a migration of capability...

...Internet now includes corporations functioning in multi-lingual environments. The user profiling method has the ability to define words in a local context on an individual basis. This includes the ability to work in foreign and even multiple languages including symbolic languages. The data access and...

9/3, K/24 (Item 9 from file: 349)  
DIALOG(R) File 349:PCT FULLTEXT  
(c) 2005 WIPO/Univentio. All rts. reserv.

00840054 \*\*Image available\*\*

**PERSONALIZED COMPUTER PERIPHERAL  
PERIPHERIQUE D'ORDINATEUR PERSONNALISE**

Patent Applicant/Assignee:

INDELELINK CORPORATION, P.O. Box 9458, Salt Lake City, UT 84109-9998, US,  
US (Residence), US (Nationality)

Inventor(s):

GERPHEIDE George, 3481 South Monte Verde, Salt Lake City, UT 84109, US,

Legal Representative:

O'BRYANT David W (et al) (agent), Morriss, Bateman, O'Bryant & Compagni,  
P.C., Suite 300, 5882 South 900 East, Salt Lake City, UT 84121, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200173745 A1 20011004 (WO 0173745)

Application: WO 2001US10579 20010329 (PCT/WO US0110579)

Priority Application: US 2000192931 20000329

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ  
EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS  
LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ  
TM TR TT TZ UA UG UZ VN YU ZA ZW  
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR  
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG  
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW  
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 10766

International Patent Class: G06F-015/173

Fulltext Availability:

Detailed Description

Detailed Description

... For example,

the merchant website cannot contain any content or subject matter that is offensive. Obviously, the term "offensive" can have broad meaning. In this context, the term generally includes material that is typically referred to as pornographic, hateful, or demeaning of others. However, the standards reflected by this definition can be modified for the particular

situation.

The merchant database 120 is organized by listing associated goods...

9/3,K/46 (Item 31 from file: 349)  
DIALOG(R) File 349:PCT FULLTEXT  
(c) 2005 WIPO/Univentio. All rts. reserv.

00296823 \*\*Image available\*\*  
INPUT SYSTEM FOR TEXT RETRIEVAL  
SYSTEME D'ENTREE POUR RECHERCHER DU TEXTE  
Patent Applicant/Assignee:  
ROSSIDES Michael T,  
Inventor(s):  
ROSSIDES Michael T,  
Patent and Priority Information (Country, Number, Date):  
Patent: WO 9514974 A1 19950601  
Application: WO 94US13279 19941129 (PCT/WO US9413279)  
Priority Application: US 93158297 19931129; US 93165676 19931213  
Designated States:  
(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AM AT AU BB BG BR BY CA CH CN CZ DE DK EE ES FI GB GE HU JP KE KG KP KR  
KZ LK LR LT LU LV MD MG MN MW NL NO NZ PL PT RO RU SD SE SI SK TJ TT UA  
UZ VN KE MW SD SZ AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE BF BJ  
CF CG CI CM GA GN ML MR NE SN TD TG

Publication Language: English  
Fulltext Word Count: 7571  
Main International Patent Class: G06F-017/30  
Fulltext Availability:

Detailed Description

Detailed Description

... Next N Letters Is Below a Threshold

A function can calculate the expected value of spelling the next N letters of a name, word, or Many Definitions of Being Probably Stuck As the procedures above demonstrate, there are many ways of mathematically defining that a speaker is probably stuck. The point is not the exact definition procedure used in a guide but that a reasonable probabilistic procedure is used at all. An example...

...of names that match the abbreviation created thus far.

b. Calculates the expected information value of the next N letters in all the words specified by the last word identifier entered .

c. If the expected value is below a certain threshold, outputs a message telling the speaker that she will probably make little progress for the next N letters.

Threshold Definitions of Being Stuck

The procedures above for defining whether a speaker is stuck include threshold values. If...